IDA

INSTITUTE FOR DEFENSE ANALYSES

Cost of Unsuitability: Assessment of Trade-offs Between the Cost of Operational Unsuitability and Research, Development, Test and Evaluation (RDT&E) Costs

Tzee-Nan K. Lo, Project Leader
Harold S. Balaban
Waynard C. Devers
Christopher S. Wait
Kristen M. Guerrera

March 2008

Approved for public release;
distribution unlimited.

IDA Paper P-4330
Log: H 08-000223

This work was conducted under contract DASW01-04-C-0003, Task BD-7-2768, for the Office of the Director, Operational Test and Evaluation. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

© 2008, 2009 Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882 • (703) 845-2000.

This material may be reproduced by or for the U.S. Government pursuant to the copyright license under the clause at DFARS 252.227-7013 (NOV 95).

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-4330

Cost of Unsuitability: Assessment of Trade-offs Between the Cost of Operational Unsuitability and Research, Development, Test and Evaluation (RDT&E) Costs

Tzee-Nan K. Lo, Project Leader
Harold S. Balaban
Waynard C. Devers
Christopher S. Wait
Kristen M. Guerrera

PREFACE

The Institute for Defense Analyses (IDA) prepared this paper for the Director, Operational Test and Evaluation, under a task titled "Assessment of Trade-offs Between the Cost of Operational Unsuitability and RDT&E Costs." The objectives of this study were: (1) estimate the costs associated with a finding in Operational Test and Evaluation that a system is operationally unsuitable; and (2) estimate the extent to which such costs can be avoided by incurring added costs during the System Development and Demonstration phase. These objectives were pursued for a small, selected set of Department of Defense acquisition program case studies.

William L. Erikson, Thomas P. Frazier, and Lance M. Roark of IDA were the technical reviewers for this paper.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

TASK OBJECTIVE

Between 1984 and 2006, 36 out of the 136 systems that underwent operational test and evaluation (OT&E) were evaluated as unsuitable [1]. In light of this high proportion, in October 2006 the Director, Operational Test and Evaluation, commissioned a study from the Institute for Defense Analyses (IDA) to answer the following two questions:

- When a system is found to be operationally unsuitable, what are the associated costs?
- To what extent can such costs be avoided by addressing unsuitability issues during the System Development and Demonstration (SDD) phase?

PROJECT SCOPE

Operational suitability is a composite evaluation that considers a system's safety, interoperability, availability, maintainability, and reliability. A study of the complete cost-based trade space for addressing unsuitability issues would therefore consider the costs associated with each of these elements being unsuitable, as well as the optimal level of suitability for each element to achieve an overall desired level of suitability. In light of time and resource constraints, however, we limited ourselves to one aspect of unsuitability—substandard reliability—whose associated costs are large, readily identifiable, and calculable using validated methods. When a system is unsuitable due to substandard reliability (e.g., low mean time between maintenance, low mean time between failures, etc.) it incurs additional life-cycle cost (LCC) for maintenance personnel, replacement parts, repair, and initial spares. Given our narrowed focus, this additional LCC is what we considered to be the cost of unsuitability.

We examined the F-22, MV-22, and C-17—three major aircraft platforms that addressed substandard reliability in different programmatic phases. Both the F-22 and MV-22 demonstrated substandard reliability during their respective initial operational test and evaluation (IOT&E) periods and were subsequently evaluated as unsuitable. Both programs then attempted to improve system suitability by investing to *fix* substandard reliability through re-design, re-engineering, and retrofit of fielded units. The

C-17, by contrast, demonstrated superior reliability at its IOT&E and was evaluated as suitable. After early flight testing revealed the system to be below contract specification for reliability, the program attempted to *avoid* an unsuitable evaluation at IOT&E by investing in reliability improvements during SDD.

For each of the three aircraft we identified the additional resources (investment) devoted to improving reliability, the resulting change in reliability, and the corresponding reduction in LCC (return). We then calculated and compared the three programs' returns on investment, adjusted for life-cycle flight hours, to provide insight into the cost-based trade-offs for addressing unsuitability issues during or after SDD.

ANALYTICAL APPROACH

The analysis for each system comprised the same four steps. First, we projected the system's primary reliability metric (PRM) at maturity both with and without additional reliability investment. Second, we identified the system's additional reliability investment. Third, we estimated the reduction in the system's LCC that resulted from the investment-driven increase in reliability. Finally, we compared the reliability investment to the LCC reduction it produced.

Projecting PRMs at Maturity

Projecting the PRM at maturity both with and without additional reliability investments required that we distinguish between the reliability growth that is attributable to investment and the reliability growth that is attributable to other factors. (This distinction is required to avoid overestimating the savings provided by investments to increase reliability.) The principal "other factor" is experience accumulated from the start of testing through early fielding. We therefore modeled reliability growth as the result of two independent mechanisms:

- *Learning-driven growth*, which occurs as minor production deficiencies are corrected, engineering change proposals are instituted, and operators, maintainers, and depots gain proficiency.
- *Investment-driven growth,* which occurs as the result of specific and concerted efforts to re-design and re-engineer components and subsystems with reliability improvements.

We first projected the system's PRM at maturity from an initial, pre-investment observation under the assumption of only learning-driven growth. This gave us expected values for the PRM at all times up to system maturity if no additional reliability

investments were made. We then observed the PRM at a post-investment observation, again projecting it to maturity under the assumption of only learning-driven growth. Finally, we measured the difference between the pre-investment and post-investment projections of the PRM at maturity. The magnitude of this difference is the increase in reliability attributable to investment-driven growth.[1]

Measuring Reliability Investment

The three systems we analyzed made investments to improve reliability at varying programmatic phases. Accordingly, the sources by which we measured the cost of the reliability investments for each system also vary.

After being evaluated as unsuitable at IOT&E, the F-22 was authorized to enter full-rate production (FRP) while it simultaneously initiated investment in post-SDD fixes to the substandard reliability revealed at IOT&E. The cost of these investments is reported in budget submissions.

After being evaluated as unsuitable at its IOT&E—operational test phase IIE (OT-IIE)—the MV-22 was authorized to remain in low-rate initial production (LRIP) while SDD was extended for 5 years to improve system suitability before the FRP decision. Additional reliability investments during those 5 years were intermingled with other contract activities, thus making budget submissions an overly inclusive measure of their cost. Program office records enable us to identify just the reliability-related expenditures.

After early flight testing revealed reliability performance below contract specification, the C-17 program made additional reliability investments during SDD to increase the likelihood of a suitable evaluation at IOT&E. Those investment costs were

---

1 Learning-driven growth follows a curve of the form $\alpha(T_M/T_\alpha)^\beta$, where α is an instantaneous measurement of reliability, T_M is the cumulative operating hours at which the system reaches maturity, T_α is the cumulative operating hours that corresponds to α, and β is an estimate of the system's learning-driven growth rate ($0<\beta<1$). This specification is considered relevant until system maturity, at which point learning-driven growth is assumed to end.

As an example of projecting the PRM at maturity using this functional form, consider the F-22 case. Its pre-investment PRM measured 0.5 hrs, it had accumulated 7,870 flight hours (FH), and we estimated its learning-driven growth rate as 0.14—the average of the F-15 and F-16 rates. With these inputs, our pre-investment projection at maturity (100,000 FH) is: 0.5 hrs × $(100{,}000/7{,}870\text{ FH})^{0.14}$ = 0.71 hrs. The F-22's post-investment PRM measured 0.69 hrs, by which time it had accumulated 38,000 FH, so the post-investment projection at maturity is: 0.69 hrs × $(100{,}000/38{,}000\text{ FH})^{0.14}$ = 0.79 hrs. Thus, the increase in the F-22's PRM that is attributable to the investment is: (0.79-0.71)/0.71hrs = 11 percent.

never broken out explicitly, but rather were rolled up in the cost of the entire development contract. We used Selected Acquisition Reports, Government Accountability Office reports, and contractor input to develop a consensus estimate of them.

Reduction in LCC from Investment-Driven Reliability Growth

We calculated the reduction in LCC that results from investment-driven reliability growth by relating the value of the system's PRM to its variable LCC components: maintenance personnel, replacement parts (consumables), depot-level repairables (DLRs), and initial sparing requirements. To do this, we used analytical tools that IDA had previously developed and validated:

- Simulation model for maintenance personnel
- Consumable and DLR cost-estimating relationships
- Initial spare demand curves

We computed the system's base case LCC to be the sum of its life-cycle costs for maintenance personnel, consumables and DLRs, and initial spares, when the PRM is equal to its requirement at maturity. That is, the base case LCC corresponded to the smallest value of the PRM for which it would not be considered a contributor to system unsuitability. Exceeding the PRM requirement at maturity would result in an LCC below the base case, and the differential would represent the savings associated with enhanced reliability. Conversely, failing to meet the PRM requirement at maturity would result in an LCC above the base case, and (in the context of substandard reliability) the differential would represent the cost of unsuitability.

Return on Investment

After measuring the system's reliability investment and calculating the LCC savings that result from that investment's improvement to the PRM, we divided the former into the latter to compute the return on investment (ROI). The ROI is useful because it permits inter-program comparisons of investment effectiveness by normalizing absolute costs and savings into relative payoffs. We also computed the ROI divided by the system's life-cycle flight hours to disentangle the effectiveness of the investment from the intensity of system usage, which would affect its potential for LCC savings. Although our sample size is limited, the magnitudes of the ROIs suggest that investing to

address unsuitability issues (namely, substandard reliability) is most effectively accomplished during SDD.

SUMMARY OF RESULTS

Fixing Substandard Reliability: F-22 and MV-22

The F-22 and MV-22 programs are considered together here because both retrospectively attempted to fix the substandard reliability that contributed to the unsuitable evaluations at their respective IOT&Es.

The PRM we used for the F-22 is its mean time between unscheduled maintenance (MTBM). The requirement at maturity (100,000 FH) for the F-22's MTBM is 1.5 hours [2]. When we extrapolated the F-22's MTBM from IOT&E (0.5 hours), however, we projected that its value at maturity would be only 0.71 hours—or 53 percent below the threshold for suitability.

The PRM we used for the MV-22 is its mean flight hours between failures—logistics ($MFHBF_{log}$). The requirement at maturity (60,000 FH) for the MV-22's $MFHBF_{log}$ is 1.4 hours [3]. When we extrapolated the MV-22's $MFHBF_{log}$ from OT-IIE (0.6 hours), however, we projected that its value at maturity would be only 0.82 hours—or 42 percent below the threshold for suitability.

Given these projected PRM shortfalls, we estimated the associated cost of unsuitability to be $6.7 billion (FY 2007 constant dollars) for each system (Table S-1). That the two costs are within rounding error of each other is coincidental.

Table S-1. F-22 and MV-22 Costs of Unsuitability (FY 2007 $B)

System	PRM	Requirement at Maturity	Test	Projected PRM at Maturity from Test	PRM Shortfall	Cost of Unsuitability
F-22	MTBM	1.5 hrs	IOT&E	0.71 hrs	53%	$6.7
MV-22	$MFHBF_{log}$	1.4 hrs[a]	OT-IIE	0.82 hrs	42%	$6.7

[a] Changed to 0.9 hours in 2001.

Avoiding Substandard Reliability: C-17

The C-17 program, in contrast to the F-22 and MV-22, attempted to *avoid* an unsuitable evaluation at IOT&E by improving substandard reliability during SDD.

The PRM we used for the C-17 is its mean time between corrective maintenance (MTBMc). The requirement at maturity (100,000 FH) for the C-17's MTBMc is 0.78 hours [4]. Early flight testing that concluded in January 1993, however, indicated that several of the system's reliability metrics, including the PRM, were below their contractually specified growth curves. When we extrapolated the C-17's MTBMc from January 1993 (0.23 hours), we projected that its value at maturity would have been only 0.42 hours—or 46 percent below the threshold for suitability.

Given this projected PRM shortfall, we estimated the potential associated cost of unsuitability to be $10.4 billion (FY 2007 constant dollars) for the C-17 (Table S-2).

Table S-2. C-17 Potential Cost of Unsuitability (FY 2007 $B)

System	PRM	Requirement at Maturity	Test	Projected PRM at Maturity from Test	PRM Shortfall	Potential Cost of Unsuitability
C-17	MTBMc	0.78 hrs	January 1993	0.42 hrs	46%	$10.4

Comparing "Fixing" and "Avoiding"

Calculating the actual and potential costs of unsuitability associated with substandard reliability for the three systems provided evidence on the first question of the study. To answer the second—To what extent can such costs be avoided by addressing unsuitability issues during the SDD phase?—we compared the effectiveness of investing to improve substandard reliability during SDD (C-17) versus retrospectively after a failed IOT&E (F-22, MV-22).

For each system we computed the ROI for improving reliability, where ROI is equal to the present value (PV) of the LCC savings resulting from investment-driven reliability growth divided by the PV of the reliability investment.[2] (Thus, in PV terms, an ROI of 2.0 means an investment saves twice as much it costs.) By this measure, the F-22 and MV-22 reliability investments were effective such that they were certainly worth undertaking, but their ROIs were far lower than that of the C-17 (Table S-3).

2 The PV computation is analogous to what in a business context would be called "discounted cash flow." The Office of Management and Budget identifies the PV calculation as "the standard criterion for deciding whether a government program can be justified on economic principles" [5].

Table S-3. Returns on Reliability Investment (PV 2007 $B)

System	Pre-Investment PRM Projection	Post-Investment PRM Projection	Gross LCC Savings	Investment	ROI
F-22	0.71 hrs	0.79 hrs	0.8	0.3	2.8
MV-22	0.82 hrs	1.56 hrs	5.0	0.9	5.7
C-17	0.42 hrs	1.10 hrs	16.1	0.9	18.3

An important caveat to these ROIs, however, is that they do not account for the fact the C-17 has the most flight hours over which to realize LCC savings and the F-22 the least. We would therefore expect that equivalent investments in each system would produce the greatest returns for the C-17 and the least for the F-22 *regardless of the programmatic phase in which the investment was made*. To normalize for this effect, we divided each system's ROI by its life-cycle flight hours in millions (Table S-4).

Table S-4. Adjusted Returns on Reliability Investment

System	ROI	Life-Cycle Flight Hours (millions)	Adjusted ROI
F-22	2.8	1.19	2.3
MV-22	5.7	2.79	2.0
C-17	18.3	5.22	3.5

The adjustment for life-cycle flight hours significantly changed the magnitude of the C-17 and MV-22 ROIs, even causing the F-22 and MV-22 to switch places in the ranked ordering. Still, even the adjusted ROIs show that the C-17's strategy of investing to improve substandard reliability during SDD produced substantially greater returns than those of the F-22 or MV-22. A plausible reason for this is that the re-design of components and subsystems during SDD—when system configuration is more easily changed—produces proportionally larger increases in reliability for a given amount of investment. In addition, it may be less expensive for contractors to conduct reliability improvement projects during SDD because research and development resources—both capital and labor—are already assembled for that program.

Because our sample is limited to three aircraft platforms, these results are only suggestive of the conclusion that investing to address unsuitability issues (namely, substandard reliability) is most effectively accomplished during SDD. The total sample of major weapon acquisition programs that would be eligible for an expanded analysis, however, is itself not particularly large. There are numerous examples like the F-22 and MV-22, for which improving suitability was an *ex post* consideration, yet very few like

the C-17, for which a potential cause of unsuitability was identified and resolved early. Thus, while our results are only illustrative of the optimality of addressing unsuitability issues during SDD, it may not be feasible to generate statistical confidence to that effect.

I. INTRODUCTION

A. BACKGROUND AND OBJECTIVES

When a system is deemed unsuitable due to substandard reliability, it necessarily incurs additional life-cycle cost (LCC)[1] to avoid, fix, or accept the consequences of its unsuitability. In this study, we estimated such additional LCC for three major aircraft systems that were all subject to substandard reliability, but invested to improve their reliability during different programmatic phases. By comparing the three systems' reliability investments to the corresponding reductions in their LCCs, our results lend insight into the cost-based trade space for addressing unsuitability issues at different programmatic phases. The systems we examined were:

- F-22 Raptor (fighter)
- MV-22 Osprey (multi-function tiltrotor)
- C-17 Globemaster III (airlift)

Both the F-22 and MV-22 programs attempted to *fix* substandard reliability after being evaluated as unsuitable at Initial Operational Test and Evaluation (IOT&E). Both programs invested in the re-design, re-test, and retrofit of fielded units with reliability improvements.[2] In so doing, they bore some additional development (re-design, re-test) and production (retrofit) costs upfront, but attempted to reduce the additional downstream LCC that accompanies substandard reliability.[3] Of interest is how the LCCs associated with the choice to fix substandard reliability compare with the LCCs had both programs made no investments to improve reliability and simply accepted the additional downstream costs.

---

1 Specifically, increases in certain operations and support cost elements and initial sparing costs.

2 Despite being evaluated as unsuitable, both programs were authorized to proceed with production. The F-22 entered full-rate production, whereas the MV-22 continued with low-rate initial production.

3 There is a second production cost associated with fixing substandard reliability—that of altering production processes and equipment to incorporate the fixes on future aircraft. But as this cost is rolled up with the entire procurement cost, we were not able to measure it separately and thus were forced to exclude it from our analyses. Still, given even generous estimates of what these costs might have been, excluding them would not materially alter our conclusions.

The C-17 program attempted to *avoid* an unsuitable evaluation at IOT&E by investing to improve reliability performance that early in System Development and Demonstration (SDD) had been below contract specification. In so doing, the program bore some additional development costs upfront (re-design, re-test), but avoided the additional downstream LCC that accompanies substandard reliability. Again, of interest is how the LCC associated with the choice to avoid potential unsuitability compares with the LCC had the program instead made no investments to improve reliability and simply accepted the additional downstream costs.

The structure of the paper is as follows: The three case studies, presented respectively in chapters II-IV, follow the same general template. We begin each case study with a brief history of the system up to the time its substandard reliability was addressed, followed by an explanation of our selection of the system's primary reliability metric (PRM), whose value at maturity we use as a proxy for overall system reliability. Next, we report the investment costs that were incurred to improve reliability (reflected in the PRM). We then detail our estimates of the LCCs associated with different values of the PRM at maturity. The LCC differential between the pre-investment projection of the PRM at maturity and the system requirement is what we call the "cost of unsuitability." The LCC differential between the pre-investment and post-investment projections of the PRM at maturity is the gross savings associated with the reliability investment. Finally, we compare those gross savings with the corresponding investment costs to calculate the return on investment (ROI). In Chapter V, we recapitulate and compare the results of our three case studies on substandard reliability to draw illustrative conclusions about the efficacy of addressing unsuitability issues at different programmatic phases.

B. GROUND RULES AND ASSUMPTIONS

Before turning to the case studies, we lay out the four ground rules and assumptions on which the study is based. Some of these have already been alluded to above, but here we give them an explicit treatment.

First, we considered only the additional development, procurement, and operations and support costs (O&S) that result from substandard reliability. There are various other costs that may be associated with unsuitability that we did not address (e.g., substandard safety or interoperability). Most significantly, perhaps, we did not address the "capability" costs of unsuitability. For example, a finding that a system is unsuitable may delay the delivery of a given capability to the warfighter, or result in the delivery of a capability whose operational availability is limited. Depending on the urgency of the need for the

capability—a function of the threat environment and, if applicable, the viability of a legacy system—the capability costs of unsuitability will vary considerably, and in some cases may be considered more important than the costs measured in this study. Future studies may endeavor to develop analytical tools that describe the complete range of unsuitability costs, but in considering only those costs associated with substandard reliability, we selected the most readily identifiable and calculable.

Second, we did not consider the option of additional procurement as a strategy for addressing substandard reliability. Although the purchase of additional units would mitigate the decreased operational availability associated with substandard reliability, it is typically not pursued for understandable reasons. It is discouraged both by the added cost (procurement and O&S), as well as by the difficulty inherent in justifying purchasing more of a system that has demonstrated substandard reliability. Moreover, even if additional procurement were found to be optimal in a given instance, doing so would set a bad precedent by rewarding a contractor despite having produced an unreliable system.

Third, the basis for the LCC differentials is the difference in the PRM before and after one specific subset of investments. That is, the basis for the F-22 and MV-22's LCC differentials is the difference in their PRMs between IOT&E (where they were evaluated as unsuitable) and a subsequent follow-on test period. Thus, we considered only the investments that contributed to the PRM difference between those test periods. Both the F-22 and MV-22 have additional reliability-related investments budgeted through at least 2009, but those investments were not considered in our analyses because their effects on actual performance have not been demonstrated. Similarly, the basis for the C-17's LCC differentials is the difference in its PRM between early flight testing that concluded in January 1993 and IOT&E. As such, we considered only the investments that contributed to the PRM difference between those test periods. All other reliability-related investments in the C-17 were omitted from our analyses.

Fourth, we did not attribute the entire difference between the pre-investment and post-investment values of the PRM to the reliability investments. The purpose of this study is to compare the cost of reliability investment to the LCC savings it produces. So as not to overstate the effect of reliability investment on LCC savings, we had to consider only the change in the PRM that resulted from investment (investment-driven growth); this entailed deducting the change in the PRM that resulted from learning (learning-driven growth). Investment-driven growth is the improvement in reliability attributable to the specific and concerted effort to re-design and re-engineer components and subsystems with reliability improvements. Learning-driven growth is the improvement in reliability

attributable to the correction of minor production deficiencies, the implementation of engineering change proposals, and the organic increase in proficiency among operators, maintainers, and depots over time. Conflating the two types of growth would be unimportant if we were analyzing mature systems because learning-driven growth is assumed to level off at system maturity. It is not negligible, however, in the early test and deployment phases from which our case study observations were drawn.

Reliability growth through system maturity is typically modeled with a curve of the form $\alpha(T_M/T_\alpha)^\beta$, where α is an instantaneous measurement of reliability, T_M is the cumulative operating hours at which the system reaches maturity, T_α is the cumulative operating hours that corresponds to α, and β is an estimate of the system's learning-driven growth rate ($0<\beta<1$).[4] This specification is considered relevant until system maturity, at which point learning-driven growth is assumed to end.

In many applications of such curves, the β parameter is interpreted as describing total reliability potential, combining the effects of learning-driven growth and investment-driven growth. That interpretation of the β parameter is not suitable for this study because, as mentioned above, it would tend to overstate the effect of investment on reliability improvement and, by extension, LCC savings. We therefore constructed reliability growth curves in which the β parameter expressed only the learning-driven growth rate, while we modeled investment-driven growth as an increase in the α parameter, i.e., as a displacement of the reliability growth curve. Figure 1 illustrates the growth effects of learning and investment for the generic reliability metric "mean time between X" (MTBX). Note that the vertical line at 100,000 operating hours represents system maturity, at which time learning-driven growth is assumed to end.

---

4 This relationship was first noted by J.T. Duane in 1964 [6] and has been adopted as the military's standard approach for projecting reliability metrics into the future [7–9].

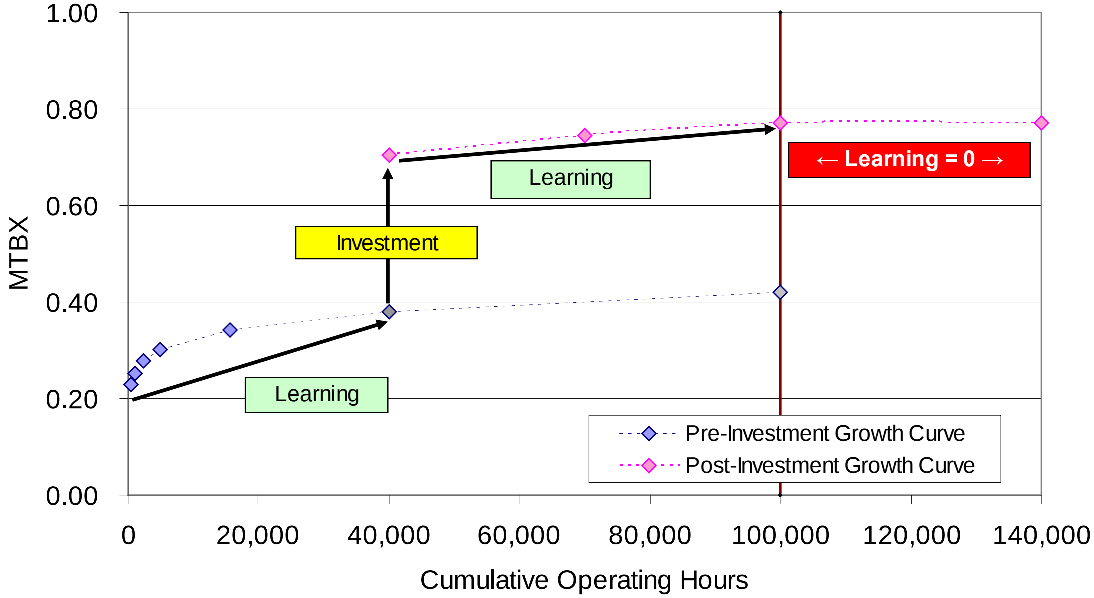

Figure 1. Learning-Driven and Investment-Driven Growth Mechanisms

II. F-22 RAPTOR

A. SYSTEM BACKGROUND

The F-22 succeeds the F-15 as the Air Force's air superiority aircraft. Its distinguishing characteristics include low observability, excellent maneuverability enhanced by thrust vectoring supercruise—a condition where supersonic flight can be sustained without using afterburner—and an advanced integrated avionics suite.

The F-22 spent over 2 decades in development—the Concept Definition phase began in 1981 and the Engineering and Manufacturing Development (EMD) phase ended in 2006. During this time the program underwent several restructurings. Most significant of these were: (1) schedule stretches early in EMD; (2) the suspended development of a two-seat model (1996); (3) the extension of EMD by an additional 9 months as recommended by an ad hoc task force (1996); and (4) the gradual reduction of the production quantity from 750 in 1990 to 175 in 2006. With the exception of the early schedule stretches, these restructurings were motivated by program cost overruns and shifting budget priorities.

The F-22 underwent IOT&E from April-September 2004. The report of the Director, Operational Test and Evaluation (DOT&E) following IOT&E judged the F-22 to be operationally effective but not suitable. Although the F-22 had already met the Operational Requirements Document (ORD) thresholds for some of its suitability metrics, it was below target for many others, including system reliability metrics. In the 3 years following IOT&E, a series of investments were made to improve the F-22's reliability, after which it tested significantly better during phase two of Follow-on Test and Evaluation (FOT&E II), conducted from June-August 2007.

B. MEAN TIME BETWEEN MAINTENANCE

The F-22's 1987 ORD establishes as a key performance parameter (KPP) that mean time between maintenance (MTBM) at maturity, i.e., 100,000 flight hours, be no less than 3 hours. The 1991 ORD update affirms this requirement and it remains the official

MTBM threshold. The importance that the ORD accords to MTBM prompted us to treat it as the PRM from which we estimated remaining variable LCC for the F-22.

The ORD requirement for MTBM is extremely ambitious—the highest MTBM previously realized among fighter/attack aircraft was only 0.62 hours (F/A-18F). In practice, the ORD requirement has been viewed more as a developmental goal than a realistic operational standard. The Air Force and Lockheed Martin have stated that 1.5 hours is the true threshold requirement for the F-22's MTBM at maturity [2], and accordingly we treated 1.5 hours as the threshold for the F-22 being suitable with respect to reliability. Thus, if the F-22 were to achieve an MTBM of 1.5 hours, then we would consider the associated cost of unsuitability to be zero. If it were to achieve an MTBM short of 1.5 hours, the additional LCC it would incur is what we would count as its cost of unsuitability.

Although but half of the ORD requirement, the Air Force MTBM requirement remains extremely ambitious. Extrapolating the pre-investment MTBM demonstrated at IOT&E (0.5 hours) according to the F-15/F-16 learning-driven growth rate (0.14) yielded a projection at maturity of 0.71 hours—or 53 percent below its requirement.[5] Even extrapolating the post-investment MTBM demonstrated at FOT&E II (0.69 hours) yielded a projection at maturity of only 0.79 hours—still 47 percent below its requirement. Between IOT&E and FOT&E II, therefore, the F-22 experienced an 11-percent displacement of its reliability growth curve. Nevertheless, to achieve an MTBM of 1.5 hours at maturity would require a 90-percent displacement of its reliability growth curve from FOT&E II.

C. THE COST OF FIXING SUBSTANDARD RELIABILITY

The cost of fixing the F-22's substandard reliability refers to the cost of all the investments between IOT&E (2004) and FOT&E II (2007)—both Research, Development, Test and Evaluation (RDT&E) and procurement expenditures—that contributed to improving MTBM from 0.5 hours to 0.69 hours. Specifically, these investments were: (1) the Reliability and Maintainability Maturation Program (RAMMP); (2) the portion of the Aircraft Engine Component Improvement Program (CIP) allocated to the F-22's F119-PW-100 engine; and (3) the implementation of RAMMP and CIP design improvements to existing aircraft via retrofitting.

---

5 We used the learning-driven growth rate for the Air Force's previous-generation fighter aircraft as the best estimate for what the F-22 might expect.

1. RAMMP

The Air Force initiated RAMMP in 2005 as a six-year research and development effort to engineer fixes to recurring failures experienced during early operational test and evaluation. Table 1 presents the RAMMP budget between IOT&E and FOT&E II.

Table 1. F-22 RAMMP Budget (FY 2007 $M)

FY 2006	FY 2007	Total
14.1	25.5	39.6

Source: FY 2008 President's Budget.

2. CIP

CIP functions as a pre-planned *ex post* suitability investment in the engines for many of the Air Force's major aircraft platforms. It funds engineering improvements in flight safety, environmental adaptability, affordability, maintainability, and reliability. As such, at least a portion of the CIP budget that was allotted to the F-22's F119-PW-100 engine between IOT&E and FOT&E II ought to be considered reliability investment. Given, however, the difficulty in determining the content of specific CIP projects, we conservatively counted the entire F119-PW-100 CIP budget between IOT&E and FOT&E II (Table 2).

Table 2. F119-PW-100 CIP Budget (FY 2007 $M)

FY 2005	FY 2006	FY 2007	Total
53.4	53.1	50.4	156.9

Source: FY 2006 President's Budget.

3. Retrofit

The design improvements that resulted from RAMMP and CIP still required additional procurement expenditures to be implemented on the existing aircraft. The six relevant program elements for this retrofitting were F22000, F22004, F22006, F22013, F22014, and F22015. Table 3 presents the combined budget for these six program elements between IOT&E and FOT&E II.

Table 3. F-22 Retrofit Budget (FY 2007 $M)

FY 2005	FY 2006	FY 2007	Total
2.0	38.3	57.9	98.3

Source: FY 2008 President's Budget.

Summing the budget streams in Tables 1 through 3 yields a total investment in reliability of \$39.6 million + \$156.9 million + \$98.3 million ≈ \$295 million (FY 2007 constant dollars). This is the cost we associated with the investment-driven growth in MTBM between IOT&E and FOT&E II.

D. LCC DIFFERENTIALS

IDA has previously developed and validated tools that use reliability metrics to estimate both variable O&S costs and initial sparing requirements. We adapted and applied those tools to the following possible outcomes for the F-22's MTBM at maturity: the projection from IOT&E (0.71 hours); the projection from FOT&E II (0.79 hours); and the Air Force's stated threshold (1.5 hours).

1. O&S Cost of Substandard Reliability

The analytical procedures developed in IDA Paper P-4134 [10] show that certain O&S cost elements vary according to MTBM, while others are either largely fixed or dependent on other parameters. Those variable O&S cost elements are (1) maintenance personnel and (2) consumables and depot-level repairables (DLRs). We derived the F-22's cost of maintenance personnel from the IDA Maintenance Estimation and Sortie Utilization Rate Evaluation (IMEASURE) model. The IMEASURE simulation estimated the number of maintenance personnel (MP) as a function of MTBM, given the system's required operational availability [10]. We derived the F-22's costs of consumables and DLRs from cost estimating relationships (CERs) with its nearest predecessor, F-15C/D/E, scaled according to relative reliability and complexity levels

Our models calculated the F-22's variable maintenance personnel and consumable/DLR costs at the squadron level for 1 year, where a squadron is composed of 24 primary aircraft authorization (PAA). Table 4 presents these results.

Table 4. F-22 Annual Variable O&S Costs (FY 2007 $M)

MTBM	0.71	0.79	1.5
Maintenance Personnel	38.0	37.2	32.9
Consumables	7.1	6.4	3.4
DLRs	51.7	46.6	24.5
Total per Squadron (24 PAA)	96.8	90.2	60.8
Total per PAA	4.0	3.8	2.5

We transformed the annual variable O&S costs per PAA into fleet-wide, life-cycle costs by multiplying by 148 PAA over the course of the F-22's 24-year service life. Table 5 reports this calculation for each of the three values of MTBM considered, as well as the LCC differentials from the requirement (1.5 hours). As alluded to in section B of this chapter, the LCC differentials represent the F-22's O&S costs of unsuitability associated with substandard reliability.

Table 5. F-22 Life-Cycle Variable O&S Costs (FY 2007 $B)

MTBM	0.71	0.79	1.5
Life-Cycle Variable O&S Cost	14.3	13.4	9.0
Life-Cycle O&S Differential	5.3	4.4	—

a. Maintenance Personnel

Given F-22 maintenance data at the two-digit work unit code level through January 2006 and the required sortie-generation rate, the IMEASURE model yielded the following equation for an F-22 squadron:

$$MP = -130.05\ln(MTBM) + 608.97$$

We multiplied MP by the average fully burdened composite rate for an F-22 maintenance crew member to obtain the annual cost of the F-22's maintenance personnel. We estimated the average fully burdened rate to be $58,128 (FY 2007 constant dollars).[6]

6 The Air Force planning factors guide AFI 65-503 lists the F-22's maintenance crew distribution by officers (0.5 percent) and enlisted members (99.5 percent). We did not have the exact distribution of pay grades, but aircraft maintenance crew officers are generally O-3s or O-4s and enlisted members E-4s or E-5s. As a middle-of-the-road estimate, we assumed that an F-22 maintenance crew is composed of 0.25 percent O-3s, 0.25 percent O-4s, 49.75 percent E-4s, and 49.75 percent E-5s. $58,128 is the summed product of these percentages and their corresponding annual composite rates [11].

b. Consumables and DLRs

We related the F-22's cost of consumables and DLRs ($\$Con_{F\text{-}22}$ and $\$DLR_{F\text{-}22}$) to those of the F-15C/D/E ($\$Con_{F\text{-}15}$ and $\$DLR_{F\text{-}15}$) using the CERs below, which come from IDA Paper P-4134 [10]:

$$\$Con_{F\text{-}22} = \$Con_{F\text{-}15} \times (MTBR_{F\text{-}15}/MTBR_{F\text{-}22}) \times (Cost_{F\text{-}22}/Cost_{F\text{-}15})$$

$$\$DLR_{F\text{-}22} = \$DLR_{F\text{-}15} \times (MTBR_{F\text{-}15}/MTBR_{F\text{-}22}) \times (Cost_{F\text{-}22}/Cost_{F\text{-}15})^{\beta}$$

$MTBR_{F\text{-}15}$ and $MTBR_{F\text{-}22}$ are the mean times between removals for the two systems. We substituted $MTBR_{F\text{-}22}$ with 4.47 × MTBM to express the CERs in terms of the PRM; this is their empirically observed relationship as documented in IDA Paper P-4134. $Cost_{F\text{-}22}$ and $Cost_{F\text{-}15}$ serve as proxies for the complexity of the two systems, and are equal to the unit recurring flyaway cost per pound of empty aircraft weight.[7] The exponent β is the elasticity of complexity that varies according to the type of DLR—that is, airframe, avionics, or propulsion DLRs.[8] The fitted β's for these three categories are, respectively, 0.71, 0.55, and 0.92; their weighted average is 0.64. This means that if the unit recurring flyaway cost per pound of the F-22 were twice that of the F-15, then the DLR costs for the F-22 would be $2^{0.64}$ (1.56) times that of the F-15.

2. Initial Sparing Cost of Substandard Reliability

The analytical procedures developed in IDA Paper P-4029 [12] show that the total initial sparing cost as a percentage of total recurring flyaway cost—called the initial sparing cost percentage (ISCP)—can be derived from MTBM. Specifically, the analysis in IDA Paper P-4029 computes that when the F-22's MTBM is 1.02 hours, its ISCP is 12.9 percent. Here we scaled those results so that when the F-22's MTBM is 0.71, 0.79, and 1.5 hours, its ISCP is 17.4 percent, 16.8 percent, and 11.2 percent, respectively.

To obtain the LCC of initial sparing for the F-22, we multiplied ISCP by the recurring flyaway cost for the entirety of the procurement period (1999-2009), lagged 1 year because the costs for initial spares were assumed to be incurred the year following

---

7 Unit recurring flyaway cost is less than the average production unit cost (APUC). It includes only flight hardware, software, system engineering, program management, and engineering change proposals. It excludes initial spares, interim logistics support, and other non-recurring start-up costs (e.g. equipment, publications) that are included in APUC.

8 For consumables, β ≈ 1 for each of the three types, and is therefore trivially omitted from the CER.

air vehicle procurement. The planned procurement quantity as of December 2006 is 175 production aircraft [13]. We used the recurring flyaway cost from IDA Paper P-4029 [12] to estimate a unit recurring flyaway cost of $123 million (FY 2007 constant dollars) that reflected the current procurement quantity. Thus, we calculated the F-22's recurring flyaway cost to be $21.5 billion (FY 2007 constant dollars).

Table 6 reports the initial sparing costs for the three values of MTBM considered. As with the O&S costs, the LCC differentials represent the F-22's initial sparing cost of unsuitability associated with substandard reliability.

Table 6. F-22 Initial Sparing Costs (FY 2007 $B)

MTBM	**0.71**	**0.79**	**1.5**
ISCP	17.4%	16.8%	11.2%
LCC of Initial Sparing	3.8	3.6	2.4
LCC Differential	1.4	1.2	—

Combining the LCC differentials from Tables 5 and 6 gives an estimate of the F-22's cost of unsuitability that is associated with substandard reliability. If the system were to achieve an MTBM of 1.5 hours at maturity, then we would consider it to be suitable with respect to reliability, and therefore its associated cost of unsuitability would be zero. If, however, it were to achieve only its extrapolated value from IOT&E of 0.71 hours, then the associated cost of unsuitability would be $5.3 billion (O&S) + $1.4 billion (initial sparing) = $6.7 billion (FY 2007 constant dollars). Similarly, if it were to achieve only its extrapolated value from FOT&E II of 0.79 hours, then the associated cost of unsuitability would be $4.4 billion (O&S) + $1.2 billion (initial sparing) = $5.6 billion. Interpolating between these three points, we constructed the F-22's LCC differentials curve that is displayed in Figure 2.

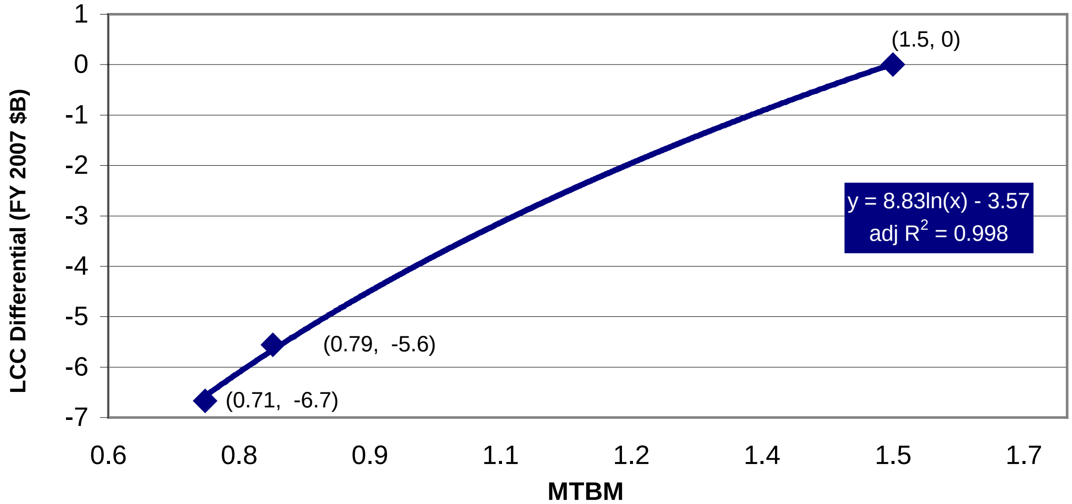

Figure 2. F-22 LCC Differentials Curve

E. RETURN ON INVESTMENT

The ROI for improving the F-22's MTBM at maturity from 0.71 hours (IOT&E) to 0.79 hours (FOT&E II) is equal to the LCC differential divided by the reliability investment between the two test periods. From the foregoing FY 2007 constant dollar calculations, the ROI is equal to ($6.7 billion – $5.6 billion)/$295 million, or 3.8. We also divided ROI by the life-cycle flight hours for the F-22 in millions (1.19), which normalized the returns with respect to program size.[9] This adjusted ROI is equal to 3.2.

A present value calculation is more appropriate, however, as it reflects the fact that while the investment cost has already been incurred, the LCC savings accrue only gradually over the course of the F-22's entire operational life. Using the 3-percent discount rate specified by the Office of Management and Budget (OMB) for FY 2007 constant dollar calculations [5], the present value ROI is equal to ($5.2 billion – $4.3 billion)/$302 million, or 2.8. Adjusted for flight hours, it is equal to 2.3. As expected, the constant dollar returns are higher than the present value returns, but both show that investment in the F-22's reliability between IOT&E and FOT&E II more than returned its cost.

As a point of interest, we also calculated the required ROI *if* the reliability-related investments through 2011 (when RAMMP concludes) were to result in an MTBM of 1.5

9 1.19 million hours ≈ 148 PAA × 24 years/PAA × 336FH/year [10].

hours. The FY 2007 present value of the LCC differential between an MTBM of 0.79 hours versus 1.5 hours is $4.3 billion; the present value of the remaining reliability investments is $456 million. Thus, the required ROI is 9.5. This means that the ROI between FOT&E II and 2011 would have to be 344 percent greater than that between IOT&E and FOT&E II.

Given that the most cost-effective investments typically occur first and then exhibit diminishing marginal returns as reliability improves, which the log-linear specification of the LCC differential curve in Figure 2 models, it seems unlikely that the F-22 will achieve an MTBM of 1.5 hours with current planned investment through 2011. Still, this is not to say that the reliability investment between FOT&E II and 2011 will not be worthwhile. The “break even” MTBM at maturity required to achieve an ROI equal to 1.0 is only about 0.85 hours, a modest 7-percent displacement of the FOT&E II reliability growth curve.

III. MV-22 OSPREY

A. SYSTEM BACKGROUND

The MV-22 succeeds the CH-46 and CH-53, and is the first rotary wing aircraft that doubles as a conventional turboprop airplane through the use of tiltrotors. The dual function of the MV-22 significantly enhances the amphibious/vertical assault capabilities of the Marine Corps via increased speed, range, and deployability.

The MV-22 program began the Concept Definition phase in 1982, EMD in 1992, low-rate initial production (LRIP) in 1996, and finally entered full-rate production (FRP) in 2005. Prior to the FRP decision, the MV-22 went through a particularly challenging developmental path as a “first of its kind” system, including: (1) threats of cancellation from 1989–1992; (2) fatal crashes in 1992 and 2000; and (3) evidence uncovered in early 2001 that some maintenance records were flawed in a way that created the appearance of enhanced maintainability.

In addition to the above, the MV-22 was evaluated as unsuitable based on IOT&E conducted from November 1999–July 2000 (OT-IIE). The fatal crash that halted OT-IIE in July 2000 may have itself been sufficient to warrant the system’s unsuitable rating, but the MV-22 also exhibited poor performance with respect to various reliability and maintainability (R&M) metrics in its Joint Operational Requirements Document (JORD). In the 5 years following OT-IIE, the Marine Corps made additional investments to improve system reliability. The MV-22 re-entered OT&E in May 2004, conducting OT-IIF from May–July 2004 and OT-IIG from March–June 2005. Only phase OT-IIG was used as the basis for its suitability evaluation, however, as the test aircraft received several significant “Block B” hardware and software upgrades in March 2005. At OT-IIG the MV-22 met all of the JORD’s R&M requirements save mean repair time for aborts.

B. MEAN FLIGHT HOURS BETWEEN FAILURES—LOGISTICS

The DOT&E report following OT-IIG states that the two key measures of reliability for the MV-22 are mean flight hours between aborts (MFHBA) and mean flight hours

between failures—logistics ($MFHBF_{log}$) [14].[10] Given that aborts may be attributable to factors unrelated to the system itself, we treated $MFHBF_{log}$ as the PRM from which we estimated remaining variable LCC for the MV-22.[11]

The original MV-22 JORD establishes that $MFHBF_{log}$ be no less than 1.4 hours at maturity (60,000 FH).[12] When OT-IIE was halted in 2000, the MV-22 had achieved an $MFHBF_{log}$ of 0.6 hours. Extrapolating this value according to the empirically observed learning-driven growth rate for the MV-22 (0.08) yielded a projection at maturity of 0.82 hours—or 42 percent below its requirement.[13] The MV-22's extended development from 2000–2005 provided an opportunity to not only meet the JORD threshold, but to achieve the contractor goal (1.5 hours) and even the program office objective (2.0 hours). During OT-IIG, the MV-22 indeed tested much better, achieving an $MFHBF_{log}$ of 1.3 hours. Extrapolating this value according to the MV-22's learning-driven growth rate yielded a projection at maturity of 1.56 hours—or 12 percent above its requirement.

Observations subsequent to OT-IIG showed continued displacements of the $MFHBF_{log}$ curve with additional investment: 1.5 hours during 2006 and 1.7 hours during the first quarter of 2007. Extrapolating these two measurements yielded projections at maturity of 1.7 hours and 1.9 hours, respectively. As the following results show, exceeding—not simply meeting—the $MFHBF_{log}$ requirement would produce significant LCC savings for the MV-22.

---

10 $MFHBF_{log}$ is numerically identical to the more common metric mean flight hours between failures (MFHBF)—the latter, in fact, appears in the MV-22's original JORD. The V-22 Blue Ribbon Panel initiated the change in 2001, suggesting that MFHBF be re-designated specifically as a *logistics reliability* metric, to distinguish it from a *mission reliability* metric (i.e., MFHBA). The rationale for having these two types of reliability metrics is that not all failures are mission-critical failures.

11 Neither MFHBA nor $MFHBF_{log}$ are KPPs in the MV-22 JORD. In fact, the MV-22 does not have any suitability-related KPPs.

12 The original JORD requirement was actually MFHBF = 1.4 hours [3]. As MFHBF and $MFHBF_{log}$ are numerically identical, however, this distinction is semantic rather than substantive. The original requirement for $MFHBF_{log}$ (as such) appears in the 2001 JORD revision and is reduced from 1.4 hours to 0.9 hours. The explanation for the reduction is that the former requirement is "outdated and inconsistent with today's configuration, technology and system complexity, measurement systems, priority, and mission profile" [15]. We nevertheless treated 1.4 hours as the threshold for reliability because it is the requirement that the system was—in theory—designed and engineered to achieve when OT-IIE was conducted.

13 We do not use the historical learning-driven growth rate for rotary aircraft systems because the MV-22 is significantly different in design and function from legacy helicopters. Rather than introduce a complexity adjustment, we simply derive the learning-driven growth rate using actual MV-22 MFHBF data from September 1998—December 2000, before the investments directed at improving reliability.

C. THE COST OF FIXING SUBSTANDARD RELIABILITY

The cost of fixing the MV-22's substandard reliability refers to the cost of the investments between OT-IIE (2000) and OT-IIG (2005) that contributed to improving $MFHBF_{log}$ from 0.6 hours to 1.3 hours. Specifically, these were the "Return to Flight" initiative and the Block A/B upgrade. Funding for such investments, however, was not directly appropriated, but rather was nested within the two more broadly defined program elements (PEs): PE 0604262N (RDT&E) and PE 059000 (retrofit). The program office helped us adjust the budget data by informing us what percentage of those appropriations were both allocable to the MV-22 Block A/B (as opposed to Block C, which was not tested at OT-IIG) and related to reliability [16]. For PE 0604262N, 55 percent of the MV-22 appropriations from 2000–2005 were both allocable to Block A/B and reliability-related appropriations. For PE 059000, 93 percent of the non-reliability/non-training system appropriations from 2000–2005 were allocable to Block A/B.

1. RDT&E

PE 0604262N was intended to correct the deficiencies in the Block A model ("Return to Flight" initiative) and make design improvements via the Block B upgrade. Two of the three line items in PE 0604262N pertained to the MV-22 (the third funded the CV-22 variant exclusively). After the 55-percent adjustment discussed above, the budget of the two MV-22 line items between OT-IIE and OT-IIG is reported in Table 7.

Table 7. MV-22 RDT&E PE 0604262N (FY 2007 $M)

Budget	FY 2000	FY 2001	FY 2002	FY 2003	FY 2004	FY 2005	Total
June 2001	30.5	—	—	—	—	—	30.5
February 2002	—	133.9	—	—	—	—	164.4
February 2003	—	—	178.3	—	—	—	342.7
February 2004	—	—	—	159.5	—	—	502.2
February 2005	—	—	—	—	149.7	—	651.9
February 2006	—	—	—	—	—	91.1	743.0

Source: FY 2002–FY 2007 President's Budgets.

2. Retrofit

PE 059000 was intended to procure retrofit kits to incorporate the RDT&E efforts of PE 0604262N. After the 93-percent adjustment discussed above, its budget between OT-IIE and OT-IIG is reported in Table 8.

Table 8. MV-22 Retrofit PE 059000 (FY 2007 $M)

Budget	FY 2001	FY 2002	FY 2003	FY 2004	FY 2005	Total
February 2002	30.7	—	—	—	—	30.7
February 2003	—	1.1	—	—	—	31.8
February 2004	—	—	–2.1[a]	—	—	29.7
February 2005	—	—	—	3.4	—	33.1
February 2006	—	—	—	—	0.4	33.5

Source: FY 2003–FY 2007 President's Budgets.
[a] Negative due to revisions from FY 2001 and 2002 appropriations.

Summing the budget streams in Tables 7 and 8 yields a total investment in reliability of \$743 million + \$33.5 million ≈ \$777 million (FY 2007 constant dollars). This is the cost we associated with the investment-driven growth in $MFHBF_{log}$ between OT-IIE and OT-IIG.

D. LCC DIFFERENTIALS

Where applicable, we adapted and applied the analytical tools from the F-22 case to the MV-22. Where necessary, we developed new CERs from legacy rotary aircraft systems. We then calculated the LCC for each of the following possible outcomes for the MV-22's $MFHBF_{log}$ at maturity: the projection from OT-IIE (0.82 hours); the JORD threshold (1.4 hours); the projection from OT-IIG (1.56 hours); and the program office's objective (2.0 hours).

1. O&S Cost of Substandard Reliability

For the MV-22, as for the F-22, the variable O&S costs with respect to reliability are (1) maintenance personnel and (2) consumables and DLRs. In lieu of an IMEASURE simulation, we created a CER between the MV-22's maintenance man hours per flight hour (MMH/FH) and $MFHBF_{log}$ to estimate its cost of maintenance personnel. We derived the MV-22's cost of consumables and DLRs from CERs using the three most recent legacy rotary aircraft (CH-46E and CH-53D/E), scaled according to relative reliability and complexity levels.

Table 9 summarizes the life-cycle variable O&S costs for the four values of $MFHBF_{log}$ considered, as well as the LCC differentials from the requirement (1.4 hours). The LCC differentials represent either the MV-22's O&S cost of unsuitability associated with substandard reliability or the O&S savings associated with enhanced reliability.

Table 9. MV-22 Life-Cycle Variable O&S Costs (FY 2007 $B)

$MFHBF_{log}$	**0.82**	**1.4**	**1.56**	**2.0**
Maintenance Personnel	4.6	3.5	3.3	2.8
Consumables + DLRs	11.9	7.0	6.2	4.9
Total Variable O&S Cost	16.5	10.5	9.5	7.7
LCC Differential	6.0	—	(1.0)	(2.8)

a. Maintenance Personnel

We defined the relationship between the MV-22's cost of maintenance personnel and system reliability by regressing MMH/FH on $MFHBF_{log}$ for the last five—more or less annual—test and evaluation periods. Using the parameters from this regression, we calculated MMH/FH for the four values of $MFHBF_{log}$ considered. We then multiplied MMH/FH by the planned life-cycle flight hours for the MV-22 (2.79 million) to get life-cycle MMH.[14] Finally, we multiplied life-cycle MMH by the average fully burdened composite hourly rate for an MV-22 maintenance crew member. We estimated the average fully burdened rate to be $72.17 (FY 2007 constant dollars).[15]

Figure 3 displays the relationship between MMH/FH on $MFHBF_{log}$ and reports the regression parameters used in the foregoing analysis.

---

14 2.79 million FH = 279 PAA × 10,000 FH per PAA ([17], [3]).

15 We did not have the exact distribution of pay grades for an MV-22 maintenance crew, but aircraft maintenance crews are approximately 2 percent officers and 98 percent enlisted members. To be consistent with the F-22 analysis, we assumed 1 percent O-3s, 1 percent O-4s, 49 percent E-4s, and 49 percent E-5s. We then multiplied these percentages by their corresponding hourly composite rates [11] to get a point estimate of $31.08 per hour for an MV-22 maintenance crew member. Our $72.17 figure is equal to $31.08 divided by 0.43, which is the observed productivity factor for the CH-53D/E from 2001-2006. This productivity factor approximates the ratio of the cost of *direct* variable maintenance personnel to the cost of *total* variable maintenance personnel for rotary aircraft systems. Productivity factors are already implicit in the IMEASURE equations, which is why we did not make a similar adjustment in the F-22 analysis.

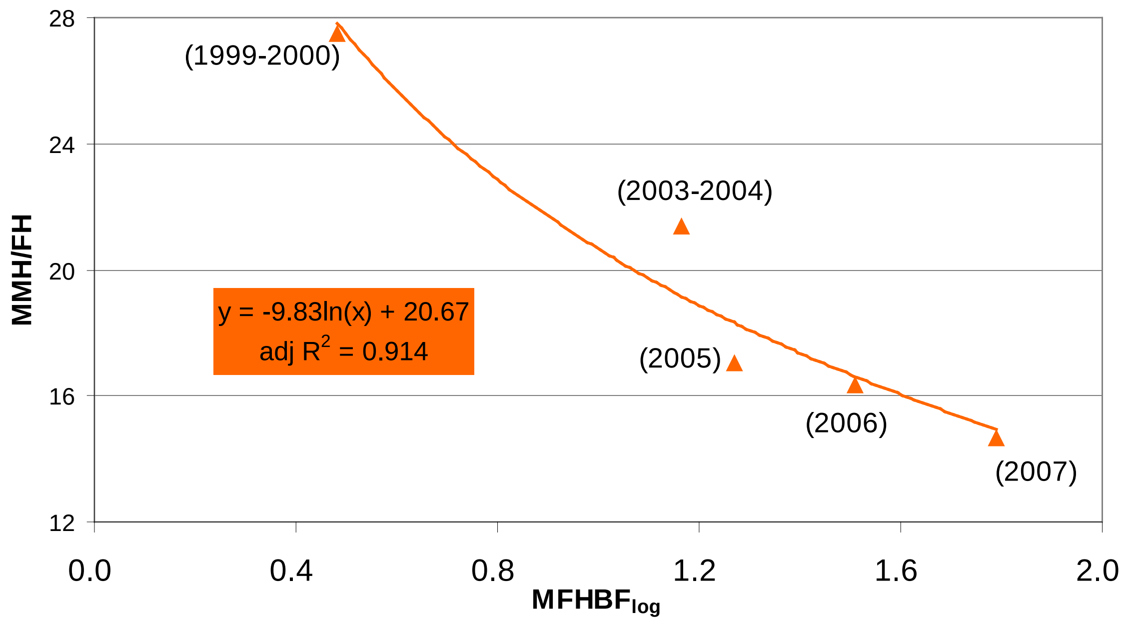

Figure 3. MV-22 MMH/FH as a Function of $MFHBF_{log}$

b. Consumables and DLRs

We calculated the LCC of MV-22 consumables and DLRs from a CER using legacy rotary aircraft systems—the CH-46E and CH-53D/E—that regressed the average cost of consumables and DLRs per non-engine maintenance action, ($Con + $DLR)/MA, on the system's unit recurring flyaway cost excluding the engine, (URFC/e). The former is a proxy for the cost of consumables and DLRs per non-engine failure; the latter, which we normalized to the 100th unit (T_{100}), is a proxy for non-engine system complexity.

We extracted the consumable costs, DLR costs, and the number and type of maintenance actions for the legacy helicopters from the Naval Aviation Logistics Data Analysis (NALDA) database. The data included all maintenance actions and all associated costs spanning from March 2005—the month in which the MV-22 entered operational testing after a 5-year layoff—to March 2007—the most current month of data at the time of this analysis. We transformed this cost data to FY 2007 constant dollars using an annual inflation factor of 2.2 percent. The unit recurring flyaway costs for the legacy helicopters, which come from contract cost data, we normalized to T_{100} given the learning curves that the contract cost data imply. We calculated the unit recurring flyaway cost for the MV-22 as the cost of T_{100} given the total buy quantity (360) [17] and the expected learning curve slope (0.91). We used a similar procedure to calculate T_{100} for the engines of the helicopters in the sample, then subtracted these values from the total

T_{100} to obtain non-engine unit recurring flyaway costs. All unit procurement costs were transformed to FY 2007 constant dollars.

We omitted from this analysis engine-related maintenance actions and engine unit recurring flyaway costs because the MV-22's AE 1107C engine is supported under a fixed-price "Power by the Hour" contract with Rolls Royce. As a commercial item acquisition, we have limited visibility into the engine-related consumable and DLR costs that Rolls Royce actually incurs in order to meet its contract requirements. In the long run, the top-level price data that are available would presumably show some partial correlation with $MFHBF_{log}$. Because the contract price is fixed in the short term, however, the cost to the government for these consumables and DLRs is not immediately sensitive to $MFHBF_{log}$. Thus, we did not consider them to be variable LCC components.

Figure 4 displays the CER between ($Con + $DLR)/MA and URFC/e, reports the regression parameters, and shows the estimated ($Con + $DLR)/MA for the MV-22 given its non-engine unit recurring flyaway cost ($63 million; FY 2007 constant dollars).

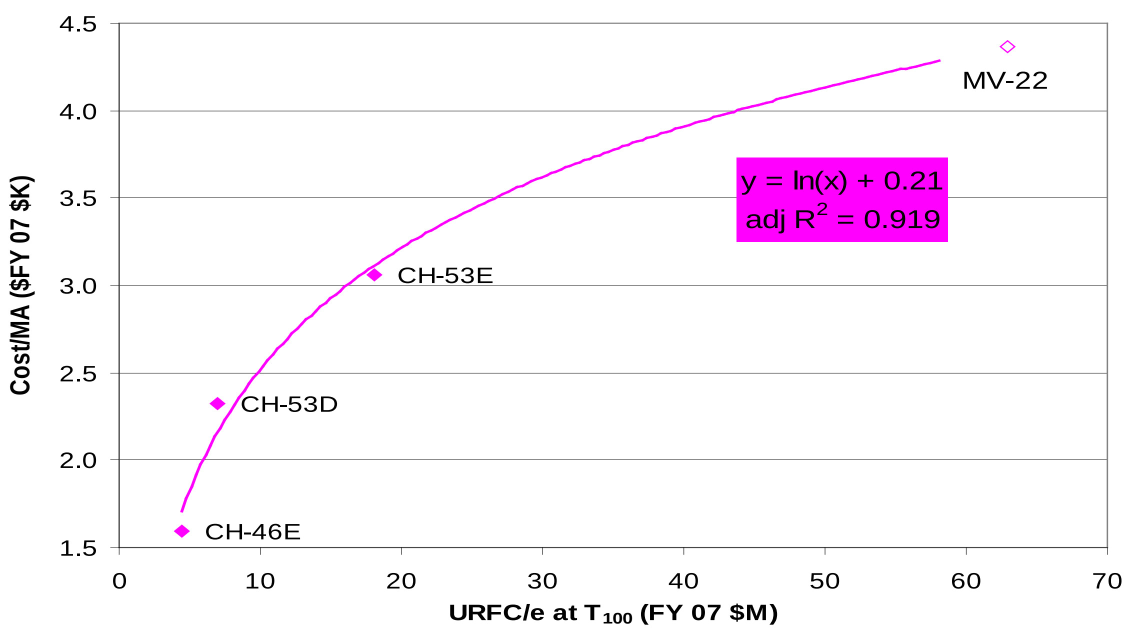

Figure 4. MV-22 Non-Engine ($Con + $DLR)/MA as a Function of System Complexity

Finally, we multiplied the MV-22's ($Con + $DLR)/MA by the expected number of life-cycle non-engine failures to calculate its LCC of consumables and DLRs for the four

values of $MFHBF_{log}$ considered.[16] The expected number of life-cycle non-engine failures is equal to the total life-cycle flight hours, divided by $MFHBF_{log}$, and multiplied by the percentage of failures that are non-engine. Consistent with NALDA data on rotary aircraft, we assumed that 20 percent of all failures are engine-related.

2. Initial Sparing Cost of Substandard Reliability

We used the same procedure for estimating the MV-22's initial sparing cost as we did for the F-22—that is, we calculated the cost of initial sparing as a percentage of total recurring flyaway cost (ISCP). Because the F-22 initial sparing analysis in Chapter II was directly traceable to the work in IDA Paper P-4029 [12], background on the ISCP equation was omitted in that section. As the parameters of the MV-22's ISCP equation are different, however, a brief explanation of the equation is in order.

The general form of the ISCP equation is: $ISCP = \alpha \times (MTBD/MTBD_0)^{\beta}$. The α parameter represents a base level ISCP for the system. The base level ISCP that we used in this analysis is the MV-22's cumulative cost of spares through 2006 ($628 million) divided by the cumulative recurring flyaway cost through 2006 ($6.7 billion), or 9.4 percent.[17] MTBD is the mean time between demand (for spares), and $MTBD_0$ is the base level MTBD that corresponds to the α parameter, i.e., the cumulative MTBD through 2006. The β parameter is the system-specific elasticity of ISCP with respect to changes in MTBD from its base level. Incidentally, the estimated β for the MV-22 is almost exactly equal to that of the F-22 β (–0.6), although the parameters we used to determine β (operating hours per day, number of aircraft, and pipeline days) are system-specific.

For this analysis, we assumed that $MFHBF_{log} = \lambda \times MTBD$ so that we could replace one for the other in the ISCP equation.[18] Accordingly, we replaced $MTBD_0$ with the cumulative value of $MFHBF_{log}$ through 2006 (1.14 hours). Given this replacement, the specific ISCP equation for the MV-22 is given by:

$$ISCP_{MV\text{-}22} = 0.094 \times (MFHBF_{log}/1.14)^{-0.6}$$

---

16 Again, maintenance actions are a proxy for failures, and we assume they exist in 1-to-1 correlation.

17 These dollar amounts are expressed in then-year dollars. We did not convert them to FY 2007 constant dollar amounts because their quotient is already index independent.

18 Note that the α parameter is unaffected by this substitution because the λ's in both numerator and denominator cancel.

To obtain the LCC of initial sparing for the MV-22, we multiplied ISCP by both the projected procurement quantity (360) and the projected non-engine unit recurring flyaway cost ($63 million; FY 2007 constant dollars). As in the analysis of consumables and DLRs, we adjusted for the fact that initial engine spares are covered under the fixed-price "Power by the Hour" contract.

Table 10 summarizes the initial sparing costs for each of the four values of $MFHBF_{log}$ considered, as well as the LCC differentials from the requirement (1.4 hours). As with the O&S costs, the LCC differentials represent either the MV-22's initial sparing cost of unsuitability associated with substandard reliability or the savings associated with enhanced reliability.

Table 10. MV-22 Initial Sparing Costs (FY 2007 $B)

$MFHBF_{log}$	**0.82**	**1.4**	**1.56**	**2.0**
ISCP	11.5%	8.3%	7.8%	6.7%
LCC of Initial Sparing	2.6	1.9	1.8	1.5
LCC Differential	0.7	—	(0.1)	(0.4)

Combining the LCC differentials from Tables 9 and 10 gives an estimate of the cost of unsuitability associated with substandard reliability for the MV-22. At the same time, given that the MV-22 is on pace to exceed its $MFHBF_{log}$ requirement, the LCC differentials show the savings from enhanced reliability.

If the MV-22 were to achieve an $MFHBF_{log}$ of 1.4 hours at maturity, then we would consider it to be suitable with respect to reliability, and therefore its associated cost of unsuitability would be zero. If, however, it were to achieve only its extrapolated value from OT-IIE of 0.82 hours, then the associated cost of unsuitability would be $6.0 billion (O&S) + $0.7 billion (initial sparing) = $6.7 billion (FY 2007 constant dollars). By contrast, if it were to achieve its extrapolated value from OT-IIG of 1.56 hours, then the savings from enhanced reliability would be $1.0 billion (O&S) + $0.1 billion (initial sparing) = $1.1 billion (FY 2007 constant dollars). And if it were to achieve the program office's objective of 2.0 hours, then the savings from enhanced reliability would be $2.8 billion (O&S) + $0.36 billion (initial sparing) = $3.2 billion (FY 2007 constant dollars). Interpolating between these four points, we constructed the MV-22's LCC differentials curve that is displayed in Figure 5.

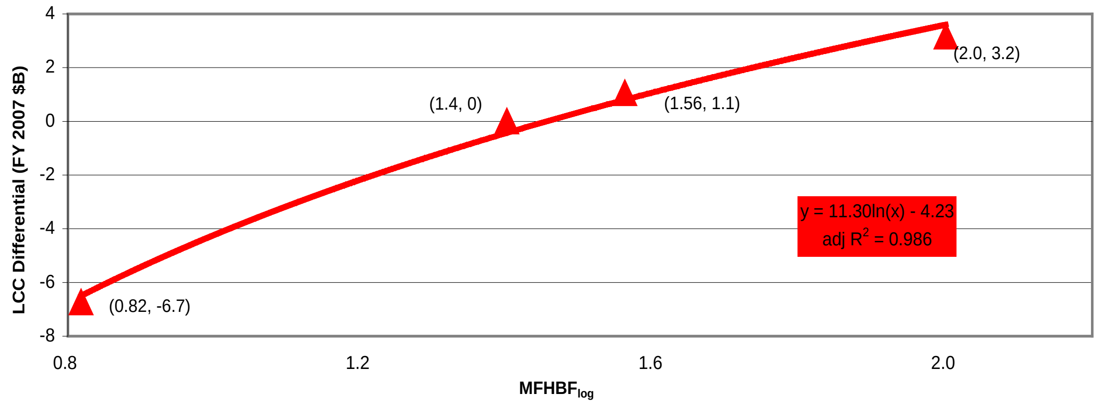

Figure 5. MV-22 LCC Differentials Curve

E. RETURN ON INVESTMENT

The ROI for improving the MV-22's projected $\text{MFHBF}_{\text{log}}$ at maturity from 0.82 hours (OT-IIE) to 1.56 hours (OT-IIG) is equal to the LCC differential divided by the reliability investment between the two test periods. From the foregoing FY 2007 constant dollar calculations, the ROI is equal to [$6.7 billion – (–$1.1 billion)]/$777 million, or 10.0. We also divided ROI by the life-cycle flight hours for the MV-22 in millions (2.79), which normalized the returns with respect to program size. This adjusted ROI is equal to 3.6.

A present value calculation again is more appropriate, however, as it reflects the fact that while the investment cost has already been incurred, the LCC savings between an $\text{MFHBF}_{\text{log}}$ of 0.82 hours versus 1.56 hours accrue only gradually over the course of the MV-22's entire operational life. Using the 3-percent discount rate specified by OMB Circular A-94 [5], the FY 2007 present value ROI is equal to [$4.3 billion – (–$0.7 billion)]/$884 million, or 5.7. Adjusted for flight hours, it is equal to 2.0. As expected, the constant dollar returns are higher than the present value returns, but both show that investment in the MV-22's reliability between OT-IIE and OT-IIG more than returned its cost.

As a point of interest, we also calculated the required ROI *if* the reliability-related investments from 2006–2009 (when the Block B retrofit is slated to conclude) were to result in achieving the program office's objective for $\text{MFHBF}_{\text{log}}$ (2.0 hours). The FY 2007 present value of the LCC differential between an $\text{MFHBF}_{\text{log}}$ of 1.56 hours versus 2.0 hours is $1.4 billion; the present value of the remaining reliability investments is $453

million. Thus, the required ROI is 3.0. This means that the ROI between OT-IIG and 2009 must only be 53 percent of that between OT-IIE and OT-IIG.

Given that the most cost-effective investments typically occur first and then exhibit diminishing marginal returns over time, it is plausible that the MV-22 will meet the program office's objective for $MFHBF_{log}$ with current planned investment through 2009. The continued improvements to $MFHBF_{log}$ observed since OT-IIG are consistent with this outcome. In any event, the "break-even" $MFHBF_{log}$ at maturity for the reliability investment from 2006–2009 is only 1.81 hours, which the projection at maturity from data in the first quarter of 2007 (1.9 hours) already exceeds.

IV. C-17 GLOBEMASTER III

A. SYSTEM BACKGROUND

The C-17 Globemaster III is the Air Force's newest and most versatile strategic airlift system, complementing the missions of both the C-5 and the C-130. The purpose of the C-17 aircraft is to modernize the airlift fleet and improve the overall capability of the military to rapidly project, reinforce, and sustain combat forces worldwide.

The Mission Element Need Statement for the C-X airlift system was approved in 1980; the Air Force selected the C-17 as the winning design in 1982; and the full-scale engineering development (FSED) contract with McDonnell Douglas began in 1985.[19] This contract placed a particular emphasis on operational suitability. It included growth curves—both required and goal—for a variety of system-level reliability metrics through maturity (100,000 FH), so that reliability performance at every test and evaluation period could be compared to a contractually specific value. Moreover, the system's reliability, maintainability, and availability performance would have to be validated prior to the Milestone IIIB decision.

In spite of these rigorous contract provisions, early flight testing that concluded in January 1993 revealed that the C-17 was below specification for some of its major reliability metrics.[20] For example, MTBR was 8 percent below its required growth curve, while mean time between corrective maintenance (MTBMc) was 46 percent below its required growth curve [4].

The January 1993 test aircraft, however, did not yet incorporate the additional reliability investments that had been initiated in 1991, the year following Secretary Cheney's decision to reduce the procurement quantity from 210 to 120. When IOT&E was conducted in 1995 following a 2-year delay to correct for a static stress failure in the

---

19 FSED corresponds to what is now called SDD. When we refer to the C-17's SDD contract, we mean its FSED contract, but use the term "SDD" to be consistent with the rest of the study.

20 Test data from this period provided by McDonnell Douglas.

wing, the C-17 tested much better as a result of these investments. The system was then on track to exceed the contractual goals for all of its major reliability metrics.

B. MEAN TIME BETWEEN CORRECTIVE MAINTENANCE

Of the major reliability metrics, MTBMc[21] was at the gravest risk for failing to meet its requirement at maturity after the January 1993 test period. Given the program's then-precarious status, substandard MTBMc might have been reason enough to delay the Milestone IIIB decision. This possibility prompted us to treat MTBMc as the PRM from which we estimated remaining variable LCC for the C-17.

The contractually specified growth curves and goal growth curves for MTBMc are reported in Table 11, where t is the cumulative number of flight hours (FH):

Table 11. C-17 MTBMc Required and Goal Growth Curves

	Required	Goal
First 1,000 FH	$0.1213t^{0.1948}$	—
1,000 – 100, 000 FH (Maturity)	$0.2151t^{0.1119}$	$0.2686t^{0.1011}$

Source: C-17 FSED Contract [4].

From Table 11, the required MTBMc as of the January 1993 test period (456 FH) is 0.40 hours, and the requirement at maturity (100,000 FH) is 0.78 hours. Yet as of January 1993, actual MTBMc was only 0.23 hours. Extrapolating this value to maturity according to the required growth curve in Table 11 yielded a projection at maturity of 0.42 hours—or 46 percent below its requirement, as mentioned in section A.

The postponement of IOT&E until 1995 provided an opportunity for the C-17 to get back to the required MTBMc by investing in additional reliability improvements. With those improvements incorporated on the test aircraft at IOT&E, the C-17 achieved an MTBMc of 0.88 hours, which at 13,500 FH already exceeded the goal at maturity. Extrapolating this value according to the required growth curve in Table 11 yielded a projection at maturity of 1.1 hours—or 41 percent above the requirement. As the following results show, exceeding—not simply meeting—the MTBMc requirement produced significant LCC savings for the C-17.

---

21 Maintenance events are grouped into three categories—those that arise from inherent failures, those that arise from induced failures, and those that arise from irreproducible failures (no-defect) events. The C-17's MTBMc metric encompasses all three categories.

C. THE COST OF AVOIDING UNSUITABILITY

The C-17's cost of avoiding unsuitability we estimate as the cost of the investments that contributed to improving MTBMc from 0.23 hours (January 1993) to 0.88 hours (IOT&E). Because such investments occurred concurrently with the SDD contract, and because the contractor was obligated to make them in order to meet contractual specifications, there was never a separate contract or program element that recorded the investment amounts. With no direct data source, therefore, we attempted to back out the C-17's additional reliability investment via the cost variance in its SDD contract from 1991–1994, as listed in the program's Selected Acquistion Reports [18].[22] Given that the investment to improve substandard reliability refers only to additional costs the contractor would not have otherwise incurred, the SDD contract cost variance should contain the entirety of these investment costs.

The SDD contract cost variance, however, also contains other elements than merely the additional reliability investment to improve suitability (e.g., normal cost growth, expanded work scope). When we explained our measurement approach to the contractor, they directed us to two elements of the contract cost variance that are definitively not related to reliability investment. First, the contractor claimed that most of the 1991 cost variance is attributable to increasing overhead costs; in response, we assumed a 70/30 split in the variance between overhead costs and reliability investment. Second, the contractor informed us that the 1994 cost variance contains $171 million in non-recurring engineering that was re-allocated to the SDD contract (from the first production contract) as part of a settlement between DoD and the contractor. This figure is confirmed by the Government Accountability Office [19].

Table 12 reports the C-17's SDD contract cost variance from 1991–1994, the overhead cost in 1991, and the settlement cost in 1994. Subtracting the latter two items from the total contract cost variance yields our estimate of the cost of the reliability investments in the C-17—$580 million (FY 2007 constant dollars). This is the cost we associated with the investment-driven growth in MTBMc from January 1993 to IOT&E. See Appendix A for a more detailed discussion of how we link this investment to growth in MTBMc between the two test events.

[22] Although the additional reliability investment began in 1991, the test aircraft during the January 1993 test period had not been retrofitted with any related improvements.

Table 12. C-17 Reliability Investment (Then-Year $M)

	1991	1992	1993	1994	Total
SDD Contract Cost Variance	221.6	110.9	91.7	352.3	776.5
Overhead	(155.1)	—	—	—	(155.1)
Settlement	—	—	—	(171.0)	(171.0)
Reliability Investment	66.5	110.9	91.7	181.3	450.4
Reliability Investment (FY 2007 $M)	89.8	145.7	117.3	227.4	580.3

D. LCC DIFFERENTIALS

Where applicable, we adapted and applied the analytical tools used in the previous cases to the C-17. Where necessary, we developed new CERs from legacy airlift systems. We then calculated the LCC for each of the following possible outcomes for the C-17's MTBMc at maturity: the projection from January 1993 (0.42 hours); the system requirement (0.78 hours); and the projection from IOT&E (1.1 hours).

1. O&S Cost of Substandard Reliability

For the C-17, as for the preceding systems, the variable O&S costs with respect to reliability are (1) maintenance personnel and (2) consumables and DLRs. We derived the C-17's cost of maintenance personnel from an application of the IMEASURE model for a comparable long-range aircraft. It estimated the number of maintenance personnel as a function of MTBMc given the C-17's required mission capable rate. We derived the C-17's cost of consumables and DLRs from a CER using C-5A/B data, scaled according to relative reliability and complexity levels.

Table 13 reports the LCC of the variable O&S costs for the three values of MTBMc considered, as well as the differentials from the base case of 0.78 hours. The differentials are the C-17's O&S cost of unsuitability given MTBMc.

Table 13. C-17 Life-Cycle Variable O&S Costs (FY 2007 $B)

MTBMc	**0.42**	**0.78**	**1.1**
Maintenance Personnel	5.9	4.8	4.2
Consumables + DLRs	28.0	21.8	15.7
Total Variable O&S Cost	33.9	26.6	19.9
LCC Differential	7.3	—	(6.7)

a. Maintenance Personnel

We could not obtain sufficient maintenance data to run a unique C-17 IMEASURE simulation, but instead used a scaled version of an IMEASURE simulation for a large, long-range aircraft like the C-17. Given the C-17's required mission capable rate, IMEASURE yielded the following equation:

$$MP/PAA = -5.99\ln(MTBMc) + 14.16$$

We multiplied MP/PAA by 174 PAA [20] over the course of the C-17's 30-year service life to obtain the fleet-wide, life-cycle number of maintenance man-years. Finally, we multiplied life-cycle maintenance man-years by the average annual fully burdened manpower rate for a C-17 maintenance crew member to obtain the system's life-cycle cost of maintenance personnel. We estimated the average fully burdened rate to be $58,754 (FY 2007 constant dollars).[23]

b. Consumables and DLRs

We related the C-17's cost of consumables and DLRs per removal ($(\$Con + \$DLR)_{C\text{-}17}/Removal$) to that of the C-5A/B with the CER below.[24] This CER is a simple transformation and combination of the CERs used in the F-22 case.[25]

$$(\$Con + \$DLR)_{C\text{-}17}/Removal = (\$Con + \$DLR)_{C\text{-}5A/B}/Removal \times (Cost_{C\text{-}17}/Cost_{C\text{-}5B})^{\beta}$$

---

23 The Air Force planning factors publication AFI 65-503 lists the C-17's maintenance crew distribution by officers (1.45 percent) and enlisted members (98.55 percent). To be consistent with the F-22 and MV-22 analyses, we assumed a pay grade distribution of 0.725 percent O-3, 0.725 percent O-4, 49.7525 percent E-4, and 49.7525 percent E-5. The $58,754 estimate is the summed product of these percentages and their corresponding annual composite rates [11].

24 Note that while we considered the consumables and DLR costs of both C-5A/B $(\$Con + \$DLR)_{C\text{-}5A/B}$, we considered only the complexity of the C-5B ($Cost_{C\text{-}5B}$). This is because the C-5A underwent a wing redesign and upgrade (1981–1987), which created a material difference between the C-5A as it was procured and the C-5A whose consumables and DLR costs we measured from 2001-2006. We chose not to drop the $(\$Con+\$DLR)_{C\text{-}5A}$ data, however, because it accounts for 31 percent of the C-5's consumables and DLR costs from 2001–2006.

25 Recall that in the F-22 CERs for consumables and DLRs, the costs are *inversely* related to the MTBR. Thus, as MTBR is equal to flight hours divided by the number of removals, the costs of consumables and DLRs are *directly* related to the number of removals. In the C-17 CER, then, we set the number of flight hours for the C-17 equal to that of the C-5, which made their number of removals comparable given their MTBRs. This, in turn, made their costs per removal comparable, as expressed in the given form of the C-17 CER for consumables and DLRs.

($\$Con + \$DLR)_{C\text{-}5A/B}$/Removal is the cost of consumables and DLRs for the C-5A/B from 2001–2006 divided by the total number of removals over that period. It is equal to \$5,564 (FY 2007 constant dollars).[26] $Cost_{C\text{-}17}/Cost_{C\text{-}5B}$ is the ratio between the unit recurring flyaway cost for the C-17 and the C-5B, which serves as a proxy for their relative levels of complexity. It is equal to 1.27.[27] The exponent β is the elasticity of complexity that varies according to the type of DLR—that is, airframe, avionics, or propulsion DLRs. The weighted average for the fitted β's for these three types is 0.70, according to the C-5A/B's specific distribution of removals from 2001–2006. Thus, as the C-17 is 1.27 times "more complex" than the C-5B, $\$DLR_{C\text{-}17}$/Removal is $1.27^{0.70}$ (1.19) times greater than $\$DLR_{C\text{-}5A/B}$/Removal. For consumables, β = 1, so $\$Con_{C\text{-}17}$/Removal is simply 1.27 times greater than $\$Con_{C\text{-}5A/B}$/Removal. From these data, we computed ($\$Con + \$DLR)_{C\text{-}17}$/Removal to be \$6,752 (FY 2007 constant dollars).

Because our CER uses the number of C-5A/B removals as reported by the Air Force's Reliability and Maintainability Information System (REMIS), our ($\$Con + \$DLR)_{C\text{-}17}$/Removal estimate of \$6,752 is implicitly the cost per *REMIS-reported* removal. For the C-17, however, we found that the annual number of removals reported in REMIS from 2001–2006 is approximately 114 percent greater than the official measurements of MTBR over that same period would imply. This disparity appears to be due to the way REMIS processes maintenance actions in general—that is, frequently double- and even triple-counting the same action. It followed that the number of C-5A/B REMIS removals was similarly overstated, which meant our initial cost per removal estimates were understated by 114 percent. We therefore multiplied our initial estimate of ($\$Con + \$DLR)_{C\text{-}17}$/Removal (\$6,752) by 2.14 (\$14,419) to correct for this understatement.

To convert cost per removal into a life-cycle cost, we multiplied \$14,419 by the expected number of removals during the C-17's operational life, where the expected number of removals is equal to life-cycle flight hours (5.22 million)[28] divided by the

---

26 \$5,564 ≈ \$2.76B/496,166 removals. The cost data comes from the Air Force Total Ownership Cost (AFTOC) database, the removals data from the Air Force's Reliability and Maintainability Information System (REMIS).

27 1.27 = \$214M/168\$M (FY 2007 constant dollars). The \$214M is the unit recurring flyaway cost of the C-17 from 1999 through 2007, and the \$168M is that of the C-5B from 1985 through 1987. We do not use the unit recurring flyaway costs over the entire procurement periods (1988–2007 and 1983–1987) because they are extremely high and unrepresentative of the majority of the procurement quantities.

28 5.22 million flight hours = 174 PAA [20] × 30,000 flight hours per PAA.

MTBR at maturity. But as our PRM is MTBMc rather than MTBR, we first derived a relationship between the two metrics so that, consistent with the other LCC components, we could express the C-17's cost of consumables and DLRs as a function of MTBMc. The empirical relationship between the F-22's MTBM and MTBR is approximately 1 to 4.5, and we attempted to derive a similar relationship using C-17 data from January 1993 to August 1995. Figure 6 displays this data and reports the regression parameters.

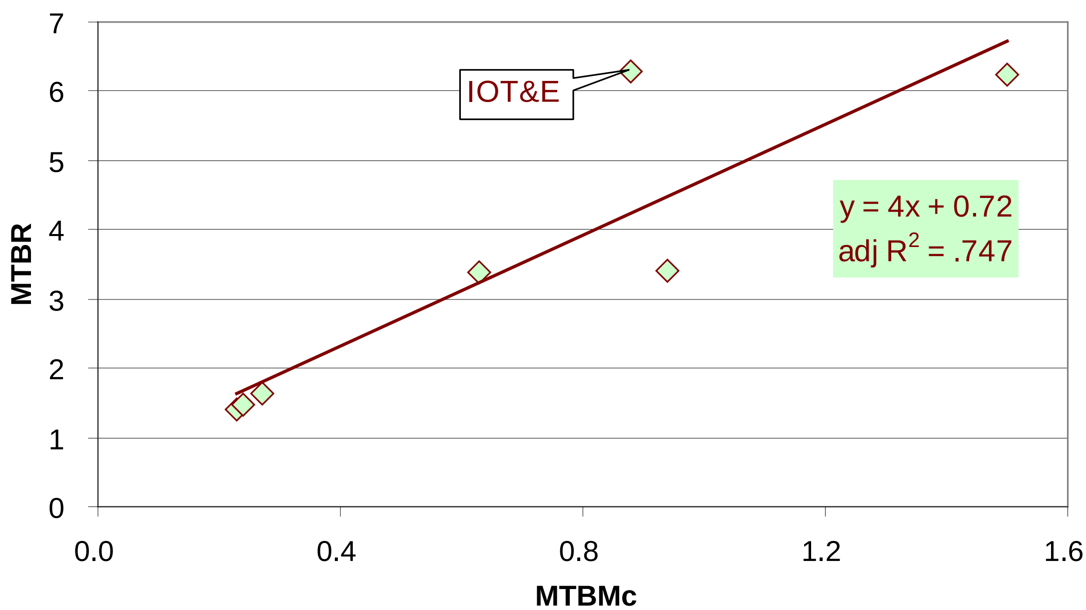

Figure 6. C-17 MTBR as a Function of MTBMc

As is apparent in Figure 6, the value of MTBR at IOT&E is an outlier from the trend line suggested by the other observations. This fact reinforces the need to derive a fixed relationship between MTBMc and MTBR as opposed to simply using the actual measured value of MTBR at IOT&E (or other test periods). Basing our LCC calculations on an atypical observation—even if it were an actual observation—would tend to misstate the savings likely to be realized from improving the C-17's reliability. We therefore ran a revised MTBMc/MTBR regression, with the IOT&E observation omitted, to generate a more consistent basis for projecting MTBR at maturity (Figure 7).

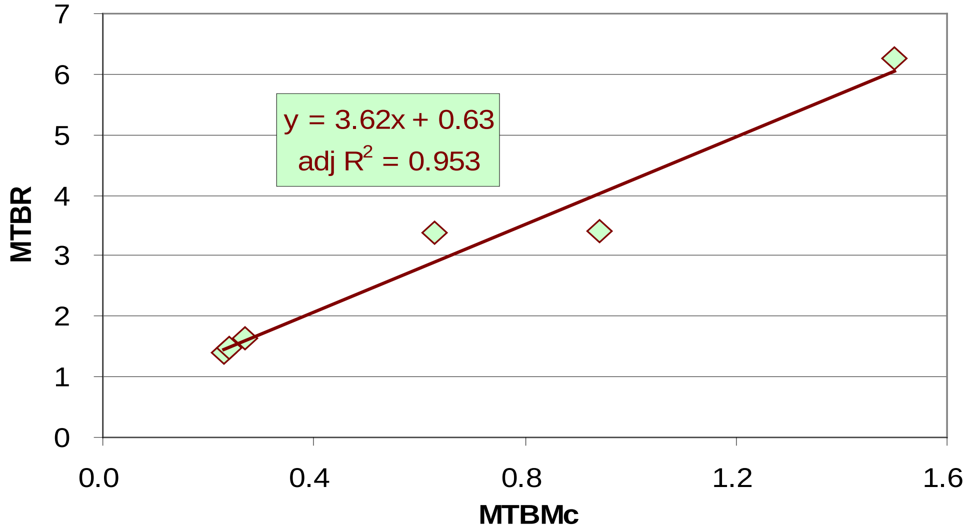

Figure 7. Revised C-17 MTBR as a Function of MTBMc

Using the revised regression parameters in Figure 7, we computed the fitted MTBRs associated with the actual MTBMc values for January 1993, IOT&E, and the system requirement. With these MTBRs, then, we were able to make projections at maturity that corresponded to the three values of MTBMc considered (Table 14).[29]

Table 14. C-17 Consumables + DLRs Summary (FY 2007 $B)

MTBMc	**0.42**	**0.78**	**1.1**
Fitted MTBR	2.69	3.45	4.78
LCC C-17 Consumables + DLRs	28.0	21.8	15.7

2. Initial Sparing Cost of Substandard Reliability

As in the previous cases, we calculated the C-17's ISCP as a function of reliability, and then multiplied ISCP by the system's total recurring flyaway cost ($40.6 billion; FY 2007 constant dollars).[30]

As mentioned in Chapter III, the ISCP equation is: ISCP = $\alpha \times (MTBD/MTBD_0)^{\beta}$. The α parameter represents a base level ISCP for the system. In this analysis, α is equal to the C-17's cumulative cost of spares from 1988–1998 ($1.6 billion) divided by the

29 The slope of the C-17's required MTBR growth curve is 0.1135 [4].

30 $40.6 billion = 190 aircraft [20] × $214 million unit recurring flyaway cost.

cumulative recurring flyaway cost over that same period ($13.4 billion), or 11.9 percent.[31] MTBD is the mean time between demand, and $MTBD_0$ is the base level MTBD that corresponds to the α parameter, i.e., the MTBD from 1998–1998. The β parameter, equal to –0.7057, is the C-17's estimated elasticity of ISCP with respect to changes in MTBD.

For this analysis, we assumed that MTBMc = γ × MTBD so that we could replace one for the other in the ISCP equation.[32] Accordingly, we replaced $MTBD_0$ with the cumulative value of MTBMc from 1988–1998, which we approximated as equal to 1.0.[33] Moreover, as dividing by 1 is trivial, this parameter simply dropped out. From this assumption, then, the specific ISCP equation for the C-17 is given by:

$$ISCP_{C\text{-}17} = 0.119 \times (MTBMc)^{-0.7057}$$

Table 15 summarizes the initial sparing costs for each of the three values of MTBMc considered. As with the O&S costs, the LCC differentials are with respect to the system requirement of 0.78 hours and represent either the C-17's initial sparing cost of unsuitability associated with the substandard reliability or the savings associated with enhanced reliability.

Table 15. C-17 Initial Sparing Costs (FY 2007 $B)

MTBMc	**0.42**	**0.78**	**1.1**
ISCP	22.0%	14.2%	11.2%
LCC of Initial Sparing	8.9	5.8	4.5
LCC Differential	3.1	—	(1.2)

Combining the LCC differentials from Tables 13 and 15 gives an estimate of the total cost of C-17's unsuitability. At the same time, given that the C-17 has demonstrated MTBMc greater than its requirement, the LCC differentials also give an estimate of the resulting LCC savings.

31 These dollar values, which come from the 1987 procurement plan, are expressed in FY 1981 constant dollars. We do not convert them to FY 2007 constant dollars because their quotient is index independent.

32 Note that this substitution does not affect the α parameter because the γ's cancel each other out.

33 This is the cumulative value of MTBMc through FOT&E, which concluded in August 1998 [21].

If the C-17 were to obtain MTBMc = 0.78 hours at maturity, then we would consider it suitable with respect to reliability, and as such it would incur no cost of unsuitability. If, however, it were to obtain only its extrapolated value from January 1993 of 0.42, then the LCC of failing to be suitable would be $7.3 billion (O&S) + $3.1 billion (initial sparing) = $10.4 billion (FY 2007 constant dollars). By contrast, if it were to obtain its extrapolated value from IOT&E of 1.1, then the life-cycle savings of enhanced suitability would be $6.7 billion (O&S) + $1.2 billion (initial sparing) = $8.0 billion. Interpolating between these three points, we constructed the C-17's LCC differentials curve displayed in Figure 8.

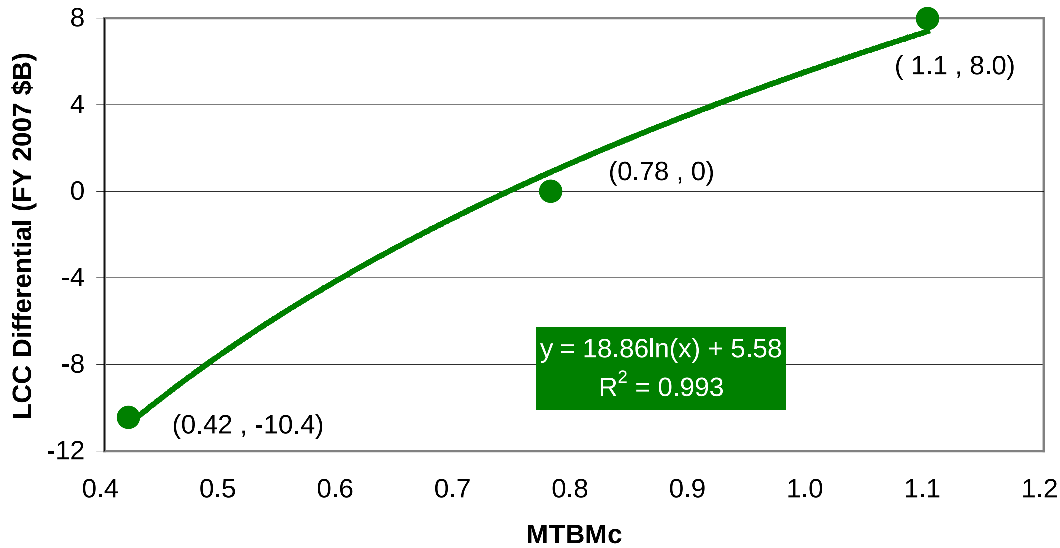

Figure 8. C-17 LCC Differentials Curve

E. RETURN ON INVESTMENT

The ROI for improving the C-17's projected MTBMc at maturity from 0.42 hours (January 1993) to 1.1 hours (IOT&E) is equal to the LCC differential divided by the reliability investment between the two test periods. From the foregoing FY 2007 constant dollar calculations, the ROI is equal to [$10.4 billion – (–$8.0 billion)]/$580 million, or 31.7. To control for program size, we divided ROI by the life-cycle flight hours for the C-17 in millions (5.22). Thus, the C-17's adjusted ROI ratio is 6.1.

As noted in the previous two cases, a present value calculation is more appropriate; it reflects the fact that while the investment cost has already been incurred, the LCC

savings between an MTBMc of 0.42 hours versus 1.1 hours accrue only gradually over the course of the C-17’s operational life. Using the 3-percent discount rate specified by OMB Circular A-94 [5], the FY 2007 present value ROI is equal to [$9.4 billion – ($6.7 billion)]/$883 million, or 18.3. Adjusting for flight hours, it is equal to 3.5. As expected, the constant dollar returns are higher than the present value returns, but both calculations show that investment in the C-17’s reliability from 1991–1994 more than returned its cost.

V. CONCLUSION

Our analysis of three major weapon acquisition programs—the F-22, MV-22, and C-17—provides evidence on the two questions this study was designed to address.

The first question was: "What are the costs associated with a finding that a system is unsuitable?" We focused on one element of unsuitability—substandard reliability—and computed the additional life-cycle costs it would impose for maintenance personnel, replacement parts (consumables), depot-level repairables, and initial sparing requirements. Given our study's narrowed focus, we referred to these additional costs collectively as the "cost of unsuitability," and found that for each of the three systems the cost of unsuitability is in the billions of dollars (Table 16). Again, we stress that system reliability is not a perfect proxy for system suitability, and that the costs associated with other elements of unsuitability (e.g., substandard safety or interoperability) are not included in this analysis.

Table 16. Cost of Unsuitability (FY 2007 $B)

System	Initial PRM Projection	PRM Requirement	Reliability Shortfall	Cost of Unsuitability
F-22	0.71 hrs	1.5 hrs	53%	$6.7
MV-22	0.82 hrs	1.4 hrs	42%	$6.7
C-17	0.42 hrs	0.78 hrs	46%	$10.4

The second question was: "To what extent can such costs can be avoided by addressing unsuitability issues during the SDD phase?" Consistent with the study's narrowed focus, we determined the extent to which the costs enumerated above could be avoided by investing to improve reliability during SDD. Our findings are illustrative but not conclusive. We determined that each of the three programs significantly reduced LCC by investing to improve reliability, and that the investments much more than returned their costs (Table 17). This result strongly suggests that addressing unsuitability issues can be of value at any programmatic phase.

Table 17. Returns on Reliability Investment (PV 2007 $B)

System	Gross LCC Savings	Investment	ROI	Life-Cycle Flight Hours (in millions)	Adjusted ROI
F-22	$0.8	$0.3	2.8	1.19	2.3
MV-22	$5.0	$0.8	5.7	2.79	2.0
C-17	$16.1	$0.8	18.3	5.22	3.5

The C-17, which alone among the three programs avoided the consequences of unsuitability by improving reliability during SDD, shows a substantially higher adjusted ROI than either the F-22 or MV-22. A plausible reason for this is that the re-design of components and subsystems during SDD—when system configuration is more easily changed—produces proportionally larger increases in reliability for a given amount of investment. In addition, it may be less expensive for contractors to conduct reliability improvement projects during SDD because research and development resources, both capital and labor, are already assembled for that program.

Because our sample is limited to three aircraft platforms, we have not definitively demonstrated that SDD is in general the best programmatic phase in which to address unsuitability issues. The total sample of DoD major weapon acquisition programs that would be eligible for such analysis, however, is in itself not particularly large. For while there are many examples like the F-22 and MV-22, in which suitability investment was an *ex post* consideration, very few programs are comparable to the C-17, in which unsuitability was identified and remedied early. Thus, while the results of the study are only illustrative of the optimality of suitability investment during SDD, it may not be feasible to generate statistical confidence to that effect.

APPENDIX A:
C-17 RELIABILITY GROWTH

This appendix provides a more detailed account of our analysis of the C-17 reliability and investment data. As noted in Chapter IV, we used as our index of C-17 reliability its mean time between corrective maintenance (MTBMc).

Figure A-1 displays data on MTBMc at January 1993 and then annually from 1993–2007. Our study is concerned with MTBMc from January 1993 through August 1998. January 1993 is when an early phase of developmental testing (DT) involving production aircraft ended; August 1998 marks the time at which the C-17 fleet reached system maturity and the contract specifications for MTBMc growth ended. Our period of analysis, however, is limited from January 1993 through June 1995, when initial operational test and evaluation (IOT&E) concluded. We measure the growth in MTBMc over just this period because of the study's orientation to addressing unsuitability issues during the development phase.

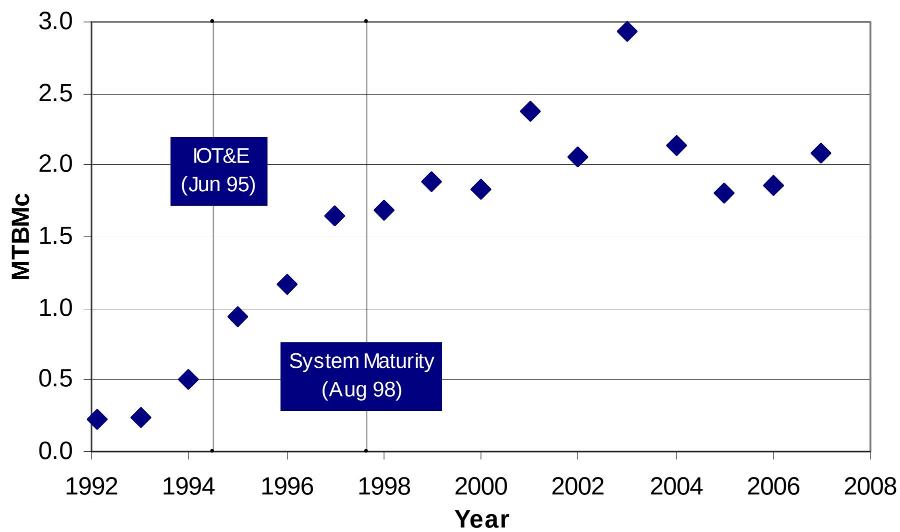

Figure A-1. C-17 Year-by-Year MTBMc (1994–2007)

As noted in Chapter IV, IOT&E was delayed until 1995 primarily to correct for a static wing stress failure, and because of the grave difficulties McDonnell Douglas

encountered in executing the C-17 contract. Concurrent with this corrective work, investments were made from 1991–1994 to improve system reliability (Table A-1). The figures in Table A-1 are estimates of annual costs incurred, that is—in contrast to budgeted amounts—estimates of actual expenditures made in those years.

Table A-1. C-17 1991–1994 Reliability Investment (FY 2007 $M)

1991	1992	1993	1994	Total
89.8	145.7	117.3	227.4	580.3

The data available to us did not indicate clearly just which aircraft were influenced by these investment expenditures. Presumably they were apportioned not only to designing reliability improvements, but also to incorporating such improvements into aircraft then in production as well as to retrofitting completed aircraft.

The crucial task of the C-17 analysis was to compute the return on these investments based on the improvement in MTBMc they produced. The major steps in our procedure were the following:

1. We began with the value of MTBMc demonstrated at the end of an early DT phase involving production aircraft (January 1993). This value was 0.23 hours; the cumulative number of flight hours as of January 1993 was 456. These 456 flight hours were accumulated on the first four production aircraft (P1-P4), which were produced from 1990–1992.
2. We extrapolated the MTBMc demonstrated at 456 hours (0.23 hours) to a value at maturity (100,000 cumulative flight hours) using the contractually required growth rate. The result was an estimated MTBMc at maturity of 0.42 hours. We interpreted the increase from 0.23 to 0.42 hours as the likely learning-driven growth in reliability that the C-17 could expect. This interpretation rests on the following:
    a. The bulk of the reliability investment was made after the 0.23 hours demonstrated in January 1993. Some of the investment from 1991 and early 1992 could have produced reliability improvements incorporated into the aircraft tested in January 1993. To the extent that this is true, our procedure understates the return on the C-17's reliability investment.

       Figure A-2 provides suggestive evidence on when the 1991–1994 investment took effect. This figure displays MTBMc observations during DT from January 1993–August 1994, initial squadron operations (ISO) from

September 1994–January 1995, and IOT&E from January–June 1995.[1] Sharp increases in MTBMc are evident between these three periods.

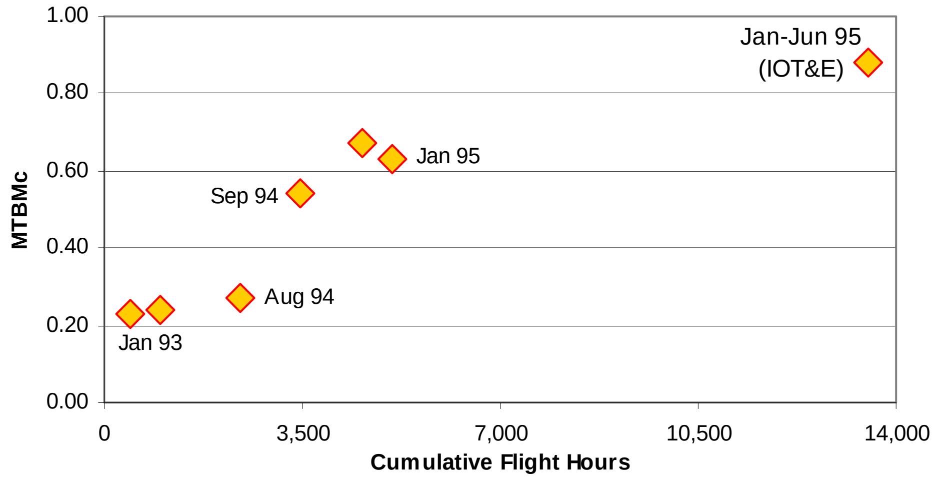

Figure A-2. C-17 MTBMc from January 1993 to June 1995

b. The C-17's full-scale engineering development (FSED) contract specifies an MTBMc of 0.1213 hours at the start of DT; a rate of increase of at least 0.1948 in the first 1,000 flight hours and 0.1119 until maturity; and, therefore, a minimum required MTBMc of 0.78 hours at maturity. The contract was complex, but the FSED portion was largely a fixed-price incentive fee arrangement. It is our understanding that the contract did not identify magnitudes for investment in improved reliability; in fact, the investment figures in Table A-1 are our estimates of the contract cost overruns that were allocated to that purpose. The expectation of the contract presumably was that the MTBMc of 0.78 hours would be achieved as minor production deficiencies were corrected, engineering change proposals were instituted, and operators and maintainers gained proficiency—which is what we define as learning-driven growth.

3. By the end of IOT&E in June 1995, the C-17 demonstrated an MTBMc of 0.88 hours and had accumulated 13,500 flight hours. Extrapolating the 0.88 hours at the contractually required growth rate of 0.1119, we estimated that the MTBMc at maturity (absent further investment) would be 1.1 hours. This measurement may not reflect the full effects of the reliability investments made from 1991–1994; to the extent that it does not, our procedure understates the return on the

1 The developmental testing data come from McDonnell Douglas and program SARs; the ISO and IOT&E data come from the C-17's Operational Reliability/Maintainability Evaluation Team.

C-17's reliability investments.[2] From the 1.1 hours we deducted our estimate of learning-driven growth (0.42 hours—step 2 above). The difference, 0.68 hours, we interpret as the investment-driven increase in MTBMc from the January 1993 observation.

Further investments intended to improve reliability were made during 1995–1998 and, as certainly would have been anticipated, MTBMc increased above our estimate of 1.1 hours. We did not attempt to compute the effect of those later investments.

---

[2] Investments to improve reliability also were made in 1995, but it is highly unlikely that these would have been incorporated into the IOT&E test aircraft given the reasonable minimum amount of time it would take to perform the retrofitting.

REFERENCES

[1] Institute for Defense Analyses/Operational Evaluation Division, "Summary of DOT&E Report to Congress," 7 October 2006.

[2] Defense Acquisition Board Briefing Chart, "F/A-22 Full Rate Production Review," 29 March 2005.

[3] "Joint Operational Requirements Document for the Joint Multi-Mission Vertical Lift Aircraft (JMVX) Change 1," 3 September 1999.

[4] C-17 FSED Contract (F33657-81-C-2108). Awarded 23 July 1982.

[5] Office of Management and Budget Circular No. A-94. Retrieved 1 March 2008 from http://www.whitehouse.gov/omb/circulars/a094/a094.html.

[6] J.T. Duane, "Learning Curve Approach to Reliability Monitoring;" *IEEE Transactions on Aerospace,* Volume 2 (1964), pp. 563–566.

[7] MIL-HDBK-781, "Department of Defense Handbook for Reliability Test Methods, Plans and Environments for Engineering, Development Qualification, and Production," 14 July 1987.

[8] MIL-HDBK-189, "Reliability Growth Management," 24 October 2000.

[9] Office of the Secretary of Defense/Cost Analysis Improvement Group, "Operating and Support Cost Estimating Guide," October 2007.

[10] Waynard C. Devers, Harold S. Balaban, Kristen M. Beal, Bruce R. Harmon, Steven E. Horne, and Neang I. Om, "F/A-22 Operating and Support Costs: Data Collection and Cost Estimation," Institute for Defense Analyses, Paper P-4134, Draft Final, June 2006.

[11] Office of the Under Secretary of Defense (Comptroller), "FY 2007 Department of Defense Military Personnel Composite Standard Pay and Reimbursement Rates," 6 April 2006.

[12] J. Richard Nelson, Bruce R. Harmon, John R. Hiller, Scot A. Arnold, Harold S. Balaban, Kristen M. Beal, John J. Cloos, Harley A. Cloud, Waynard C. Devers, Howard J. Manetti, James A. Myers, Neang I. Om, and Joseph W. Stahl, "F/A-22 Independent Cost Estimate," Institute for Defense Analyses, Paper P-4029, August 2005.

[13] Office of the Under Secretary of Defense (Acquisition, Technology and Logistics), "F-22A Selected Acquisition Report," 31 December 2006. Retrieved 1 March 2008 from the Defense Acquisition Management Information Retrieval database.

[14] Office of the Director, Operational Test and Evaluation, "V-22 Osprey Program: Report on Operational and Live Fire Test and Evaluation," September 2005.

[15] "V-22 Program Status," April 2002. Retrieved 1 March 2008 from http://www.g2mil.com/ReporttoCongress04-2002.pdf.

[16] NAVAIR V-22 Director of Logistics, "Rough Cut of R&M Investment," email dated 12 July 2007.

[17] Office of the Under Secretary of Defense (Acquisition, Technology and Logistics), "V-22 Selected Acquisition Report," 31 December 2006. Retrieved 1 March 2008 from the Defense Acquisition Management Information Retrieval database.

[18] Office of the Under Secretary of Defense (Acquisition and Technology), "C-17 Selected Acquisition Report," 31 December 1991, 31 December 1992, 30 September 1993, 31 December 1993, and 31 December 1994. Retrieved 1 March 2008 from IDA-owned compact discs containing electronic copies of the reports.

[19] GAO/T-NSIAD-94-172, "The C-17 Proposed Settlement and Program Update," 28 April 1994.

[20] Office of the Under Secretary of Defense (Acquisition, Technology and Logistics), C-17 Selected Acquisition Report, 31 December 2006. Retrieved 1 March 2008 from the Defense Acquisition Management Information Retrieval database.

[21] Air Mobility Command, "Final Report C-17 Logistics Reliability, Maintainability, and Availability Follow-on Operational Test and Evaluation of the C-17 Aircraft," April 1999.

ABBREVIATIONS

AFTOC	Air Force Total Ownership Cost
APUC	average production unit cost
B	billion
CER	cost estimating relationship
CIP	[Aircraft Engine] Component Improvement Program
Con	consumables
DLR	depot-level repairable
DoD	Department of Defense
DOT&E	Director, Operational Test and Evaluation
DT	developmental testing
EMD	Engineering & Manufacturing Development
FH	flight hours
FOT&E	Follow-on Test and Evaluation
FRP	full-rate production
FSED	full-scale engineering development
FY	fiscal year
IDA	Institute for Defense Analyses
IMEASURE	IDA Maintenance Estimation and Sortie Utilization Rate Evaluation
IOT&E	Initial Operational Test and Evaluation
ISCP	initial sparing cost percentage
JORD	Joint Operational Requirements Document
KPP	key performance parameter
LCC	life-cycle cost
LRIP	low-rate initial production
M	million
MA	maintenance action
MFHBA	mean flight hours between aborts

MFHBF	mean flight hours between failures
$MFHBF_{log}$	mean flight hours between failures—logistics
MMH/FH	maintenance man hours per flight hour
MP	maintenance personnel
MTBD	mean time between demand
MTBM	mean time between maintenance
MTBMc	mean time between corrective maintenance
MTBR	mean time between removals
MTBX	mean time between X
NALDA	Naval Aviation Logistics Data Analysis
O&S	operations and support
OMB	Office of Management and Budget
ORD	Operational Requirements Document
OT&E	operational test and evaluation
PAA	primary aircraft authorization
PE	program element
PRM	primary reliability metric
PV	present value
R&M	reliability and maintainability
RAMMP	Reliability and Maintainability Maturation Program
RDT&E	Research, Development, Test and Evaluation
REMIS	Reliability and Maintainability Information System
ROI	return on investment
SAR	Selected Acquisition Report
SDD	System Development and Demonstration
URFC/e	unit recurring flyaway cost excluding the engine

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*
xx-03-2008

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
Oct 06-Mar 08

4. TITLE AND SUBTITLE
Cost of Unsuitability: Assessment of Trade-offs Between the Cost of Operational Unsuitability and Research, Development, Test and Evaluation (RDT&E) Costs

5a. CONTRACT NUMBER
DASW01-04-C-0003

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S), Last Names First
Lo, Tzee-Nan K.
Balaban, Harold S.
Devers, Wayard C.
Wait, Christopher S.
Guerrera, Kristen M.

5d. PROJECT NUMBER

5e. TASK NUMBER
BD-7-2768

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 22311-1882

8. PERFORMING ORGANIZATION REPORT NUMBER
P-4330

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Director, Operational Test and Evaluation
1800 Defense Pentagon
Room 3D947
Washington, DC 20301

10. SPONSOR/MONITOR'S ACRONYM(S)
DOT&E

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
More than one fourth of all major weapon systems that underwent Operational Test and Evaluation from 1984-2006 were evaluated as unsuitable. A frequent and significant contributing factor in these unsuitable evaluations was substandard reliability. When a system is deemed unsuitable due to substandard reliability, it necessarily incurs additional costs to avoid, fix, or accept the consequences of its unsuitability. We estimate such costs for three aircraft systems (F-22, MV-22, and C-17) that were all subject to substandard reliability, but invested to improve it during different programmatic phases. The F-22 and MV-22 programs attempted to fix substandard reliability after being evaluated as unsuitable at Initial Operational Test and Evaluation (IOT&E). The C-17 attempted to avoid an unsuitable evaluation at IOT&E by investing to improve its substandard reliability during the System Development and Demonstration phase. Our results provide insight into the cost-based trade space for addressing unsuitability issues at different programmatic phases. We found that (1) being unsuitable due to substandard reliability is costly; (2) investments to improve reliability exhibit strong returns; and (3) earlier investment in reliability yields higher returns.

15. SUBJECT TERMS
Operational Suitability; Major Defense Acquisition Programs (MDAPs); Reliability and Maintainability (R&M); Military Aircraft; C-17 Aircraft, F-22 Aircraft, MV-22 Aircraft

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
Unclassified	Unclassified	Unclassified	Same as Report	63

19a. NAME OF RESPONSIBLE PERSON (Last Name First)
Seglie, Ernest A.

19b. TELEPONE NUMBER (*Include area code*)
703-697-7247

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI-Std Z39-18

The Institute for Defense Analyses is a non-profit corporation that administers three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.